



TESIS-SS14 2501

**ANALISIS KELOMPOK DENGAN METODE *KERNEL*
K-MEANS
(STUDI KASUS PENGELOMPOKAN
KABUPATEN/KOTA DI INDONESIA BERDASARKAN
PENDUDUK DENGAN FAKTOR-FAKTOR RISIKO
PENYEBAB PENYAKIT HIPERTENSI)**

SITI MAYSAROH
1313 201 711

DOSEN PEMBIMBING
Dr. Santi Wulan Purnami, S.Si, M.Si

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015



THESIS-SS14 2501

**CLUSTER ANALYSIS USING KERNEL
K-MEANS METHODS
(CASE STUDY GROUPING OF DISTRICT/CITY
BASED ON POPULATION WITH HYPERTENSION
RISK FACTORS)**

SITI MAYSAROH
1313 201 711

SUPERVISOR
Dr. Santi Wulan Purnami, S.Si, M.Si

PROGRAM OF MAGISTER
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

**ANALISIS KELOMPOK DENGAN METODE *KERNEL K-MEANS*
(STUDI KASUS PENGELOMPOKAN KABUPATEN/KOTA DI INDONESIA
BERDASARKAN PENDUDUK DENGAN FAKTOR-FAKTOR RISIKO
PENYEBAB PENYAKIT HIPERTENSI)**

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember
Oleh :

SITI MAYSAROH
NRP. 1313 201 711

Tanggal Ujian : 14 Juli 2015
Periode Wisuda : September 2015

Disetujui Oleh:

1. **Dr. Santi Wulan Purnami, S.Si, M.Si**
NIP. 19720923 199803 2 001

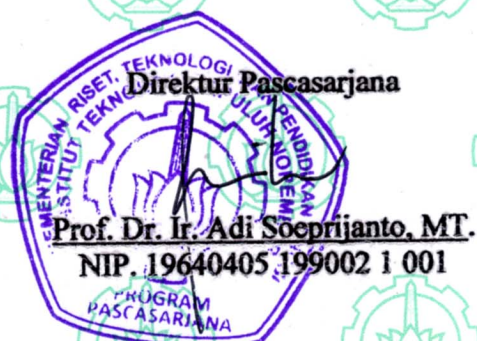
(Pembimbing)

2. **Dr. Muhammad Mashuri, MT**
NIP. 19620408 198701 1 001

(Penguji)

3. **Dr. Bambang Wijanarko Otok, M. Si**
NIP. 19681124 199412 1 001

(Penguji)



ANALISIS KELOMPOK DENGAN METODE *KERNEL K-MEANS* (STUDI KASUS PENGELOMPOKAN KABUPATEN/KOTA DI INDONESIA BERDASARKAN PENDUDUK DENGAN FAKTOR-FAKTOR RISIKO PENYEBAB PENYAKIT HIPERTENSI)

Nama mahasiswa : Siti Maysaroh
NRP : 1313 201 711
Pembimbing : Dr. Santi Wulan Purnami, S.Si, M.Si

ABSTRAK

K-means clustering adalah salah satu metode pengelompokan nonhirarki yang paling banyak digunakan karena algoritma K-means lebih sederhana dan cepat prosesnya. Namun *K-Means Clustering* memiliki kelemahan di dalam memproses data yang berdimensi banyak, yang sifatnya tidak dapat dipisahkan secara linier. Dalam perkembangannya, fungsi kernel digunakan untuk mengatasi pengelompokan data yang bersifat tidak linier. Metode pengelompokan yang menggunakan fungsi kernel tersebut dikenal dengan *Kernel K-Means Clustering*. Dalam penelitian ini akan dibahas mengenai konsep pengelompokan dengan metode Kernel K-Means, serta penerapannya pada data simulasi dan data riil. Data simulasi yang digunakan berupa data yang bersifat linier dan non linier. Sedangkan data riil yang digunakan adalah data persentase penduduk yang memiliki faktor-faktor risiko penyebab penyakit hipertensi. Tujuan dari penelitian ini adalah untuk mengkaji konsep metode kernel k-means *clustering*, membandingkan pengelompokan pada data simulasi dan data riil, serta melakukan pengelompokan kabupaten/kota di Indonesia berdasarkan faktor-faktor resiko penyebab penyakit hipertensi pada kelompok usia produktif menggunakan metode *K-Means Clustering* dan *Kernel K-Means Clustering*. Algoritma Kernel K-Means pada prinsipnya hampir sama dengan metode K-Means, perbedaannya terletak pada inputnya, yaitu pada Kernel K-Means data dipetakan pada dimensi baru menggunakan fungsi kernel. Dari hasil studi simulasi dengan kriteria kesesuaian plot, nilai *clustering error* dan nilai indeks validitas kelompok, metode pengelompokan Kernel K-Means lebih baik daripada metode K-Means. Begitu pula pada data riil hasil pengelompokan dengan metode Kernel K-Means lebih baik daripada K-Means.

Kata Kunci: Analisis Kelompok, *K-Means Clustering*, *Kernel K-Means Clustering*, indeks S_{Sbw}

CLUSTER ANALYSIS USING KERNEL K-MEANS METHODS (CASE STUDY GROUPING OF DISTRICT/CITY BASED ON POPULATION WITH HYPERTENSION RISK FACTORS)

By : Siti Maysaroh
Student Identity Number : 1313 201 711
Supervisor : Dr. Santi Wulan Purnami, S.Si, M.Si

ABSTRACT

K-means clustering is a method of nonhierarchi grouping which is most widely used due to its simple and quick process algorithm of K-means. The K-Means Clustering had weakness in processing non-linier and multi dimensions data. To overcome these problems, kernel function is used to make the grouping process or better known as Kernel K-Means Clustering. In this study will be discussed on the concept of grouping the Kernel K-Means method, and its application to the simulation data and the real data. The simulation data was data that linearly separable and non linearly separable. While the real data was data the percentage of people who have these risk factors cause hypertension. The purpose of this study is to assess the Kernel K-Means clustering, compare grouping the simulated data and real data, as well as grouping districts / cities in Indonesia based on risk factors for disease-causing hypertension in the productive age group using the K-Means Clustering and Kernel K-Means Clustering. Kernel K-Means algorithm on a principle similar to K-Means method, the difference lies in the input, namely the Kernel K-Means mapped data on a high dimension using kernel functions. Based on application to the simulation data and the real data , with the criteria for grouping that is plotting, value of clustering error and validity index, grouping method using Kernel K-Means is better than K-means.

Keyword: Cluster Analysis, *K-Means Clustering*, *Kernel K-Means Clustering*, *S_Sbw Indices*

KATA PENGANTAR

Segala puji milik Allah SWT, Dzat Yang Maha Esa, syukur Alhamdulillah penulis panjatkan kehadiran Allah SWT yang telah memberikan limpahan Rahmat sehingga penulis dapat menyelesaikan tesis dengan judul:

**“ANALISIS KELOMPOK DENGAN METODE *KERNEL K-MEANS*
(Studi Kasus Pengelompokan Kabupaten/kota di Indonesia Berdasarkan
Penduduk Dengan Faktor-Faktor Risiko Penyebab Penyakit Hipertensi) ”**

Dalam menyusun tesis ini, penulis memperoleh banyak bantuan dari berbagai pihak, baik secara langsung maupun tidak langsung, untuk itu pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Suamiku tercinta Arif Widyadarma, untuk semua pengorbanan, kesabaran, pengertian, dukungan yang luar biasa, dan keikhlasannya menggantikan peran sebagai ibu selama penulis melanjutkan studi di ITS. Anak-anakku tersayang Al Fatih Fayzul Haq, Faris Azzamul Haq, dan Farras Ihsanul Haq atas doa dan kesabaran dalam penantiannya.
2. Bapak dan Ibu, Ayah (alm) dan Ibu, atas dukungan, doa, dan kasih sayang yang tak terbalaskan. Serta semua keluarga yang telah memberikan dukungan, semangat, dan doanya.
3. Jajaran pimpinan Pusdiklat BPS, BPS D.I Yogyakarta, dan pimpinan BPS Kabupaten Kulon Progo yang telah memberi kesempatan kepada penulis untuk melanjutkan studi program S2 di ITS.
4. Ibu Dr. Santi Wulan Purnami, S.Si, M.Si selaku dosen pembimbing yang telah banyak meluangkan waktu serta dengan penuh kesabaran dan keikhlasannya dalam memberikan bimbingan, saran dan masukan serta motivasi dalam penyusunan tesis ini.
5. Bapak Dr. Muhammad Mashuri, MT, dan Bapak Dr. Bambang Widjanarko Otok, M.Si, selaku dosen penguji yang telah banyak memberikan koreksi dan saran pada penulisan tesis serta arahnya kepada penulis.

6. Bapak Dr. Muhammad Mashuri, MT selaku Ketua Jurusan Statistika FMIPA ITS Surabaya dan Bapak Dr. Suhartono, M.Sc selaku Koordinator Program Studi, dan Dr. Vita Ratnasari, S.Si, M.Si selaku Dosen Wali.
7. Bapak dan Ibu dosen pengajar serta seluruh jajaran karyawan-karyawati jurusan Statistika FMIPA ITS Surabaya.
8. Mba' Shofi dan Mas Syahrul, yang telah meluangkan waktu untuk berbagi ilmu dan atas bantuannya.
9. Teman-teman BPS angkatan 7, Ratna, Arifah, Mbak Reny, Mbak Lilis, Metty, Tika, Mbak Eta, Maya, Devy, Rini, Mas Ade, Untung, Nora, Mas Cahyo, Hadi, Gama, Choey, Aal, dan Hery, terima kasih atas segala bantuan dan kebersamaan selama menjalani pendidikan di ITS.
10. Teman-teman kantor BPS Kabupaten Kulon Progo atas bantuannya selama masa studi.
11. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu dalam penyusunan tesis ini.

Akhir kata, semoga segala kebaikan yang telah diberikan kepada penulis, mendapatkan pahala dari Allah SWT dan penulis menyadari bahwa tesis ini masih jauh dari kesempurnaan. Untuk itu, kritik dan saran yang bersifat membangun sangat penulis harapkan demi kesempurnaan tesis ini. Semoga tesis ini dapat bermanfaat bagi sesama. Aamiin Ya Robbal 'Alamin.

Surabaya, Agustus 2015

Penulis

DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN.....	i
ABSTRAK.....	iii
ABSTRACT.....	v
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	7
2.1 Analisis Cluster.....	7
2.1.1 Ukuran Kemiripan	7
2.1.2 Metode Pengelompokan	9
2.1.3 Metode <i>K-Means Clustering</i>	10
2.1.4 Metode Kernel <i>K-Means Clustering</i>	13
2.1.5 Evaluasi Hasil Pengelompokan	15
2.2 Faktor Risiko Penyakit Hipertensi	21
BAB III METODOLOGI PENELITIAN	25
3.1 Kajian Teori	25
3.2 Kajian Terapan.....	25
3.2.1 Penerapan pada Data Simulasi	26
3.2.2 Studi Kasus pada Data Riil (Data Persentase Penduduk dengan Faktor Risiko Penyebab Penyakit Hipertensi Menurut Kabupaten/Kota)	27
BAB IV HASIL DAN PEMBAHASAN	37
4.1 <i>Kernel K Means Clustering</i>	37

4.1.1	Sifat-Sifat Kernel	37
4.1.2	Algoritma Kernel K-Means <i>Clustering</i>	40
4.1.3	Kompleksitas Algoritma Kernel K Means	42
4.2	Perbandingan Pengelompokan metode K-Means dan Kernel K-Means pada Data Simulasi	43
4.3	Penerapan metode <i>K-Means</i> dan <i>Kernel K-Means</i> untuk Pengelompokan Kabupaten/Kota di Indonesia	50
4.3.1	Rekomendasi Jumlah Kelompok (<i>Cluster</i>) yang Optimal	54
4.3.2	Pengelompokan Kabupaten/Kota di Indonesia menggunakan metode KM dan KKM.....	55
4.3.3	Interpretasi Hasil Pengelompokan.....	61
BAB V KESIMPULAN DAN SARAN		73
5.1	Kesimpulan.....	73
5.2	Saran.....	74
DAFTAR PUSTAKA.....		75

DAFTAR TABEL

	Halaman
Tabel 2.1	Evaluasi Performansi 11 Ukuran Validitas 20
Tabel 3.1	Variabel dan Label yang Digunakan dalam Penelitian..... 33
Tabel 4.1	Perbedaan Rata-rata Nilai <i>Clustering Error</i> (<i>e</i>) Pengelompokan dengan Metode KM dan KKM..... 48
Tabel 4.2	Indeks Validitas Kelompok pada Metode KM dan KKM pada Data Simulasi dan Data Penelitian..... 49
Tabel 4.3	Output Koefisien Determinasi (R^2)..... 50
Tabel 4.4	Output ANOVA (Uji F)..... 50
Tabel 4.5	Ringkasan Pengujian Regresi Secara Parsial (Uji t)..... 51
Tabel 4.6	Jumlah Kabupaten menurut Provinsi di Indonesia tahun 2013.. 52
Tabel 4.7	Statistik Deskriptif Variabel Penelitian..... 53
Tabel 4.8	Indeks Validitas Kelompok untuk Menentukan jumlah Kelompok Optimal..... 54
Tabel 4.9	Nilai Pusat Kelompok hasil pengelompokan dengan KM dengan k=2..... 56
Tabel 4.10	Nilai Pusat Kelompok hasil pengelompokan dengan KM dengan k=3..... 57
Tabel 4.11	Nilai Pusat Kelompok hasil pengelompokan dengan KKM dengan k=2..... 58
Tabel 4.12	Nilai Pusat Kelompok hasil pengelompokan dengan KKM dengan k=3..... 59
Tabel 4.13	Jumlah Anggota per Kelompok pada pengelompokan dengan metode KKM dengan k=2..... 61
Tabel 4.14	Jumlah Anggota per Kelompok pada pengelompokan dengan metode KKM dengan k=3..... 63
Tabel 4.15	Jumlah Anggota per Kelompok pada pengelompokan menurut faktor kejiwaan menggunakan metode KKM dengan k=3..... 64
Tabel 4.16	Jumlah Anggota per Kelompok pada pengelompokan menurut

	faktor pola makan menggunakan metode KKM dengan k=3....	66
Tabel 4.17	Jumlah Anggota per Kelompok pada pengelompokan menurut faktor aktivitas fisik menggunakan metode KKM dengan k=3.	68
Tabel 4.18	Jumlah Anggota per Kelompok pada pengelompokan menurut faktor gaya hidup menggunakan metode KKM dengan k=3....	70
Tabel 4.19	Jumlah Anggota per Kelompok pada pengelompokan menurut faktor lain-lain menggunakan metode KKM dengan k=3.....	71

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Kerangka Konsep	28
Gambar 3.2 Alur Kerja Metode <i>K-Means Clustering</i>	34
Gambar 3.3 Alur Kerja Metode <i>Kernel K-Means Clustering</i>	35
Gambar 4.1 Perbedaan Plot data Linier Hasil Pengelompokan Metode KM dan KKM.....	45
Gambar 4.2 Perbedaan Plot data Non Linier Hasil Pengelompokan Metode KM dan KKM.....	47
Gambar 4.3 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan <i>K-Means Clustering</i> dengan jumlah Kelompok sebanyak 2 Kelompok.....	56
Gambar 4.4 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan <i>K-Means Clustering</i> dengan jumlah Kelompok sebanyak 3 Kelompok.....	57
Gambar 4.5 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan <i>Kernel K-Means Clustering</i> dengan jumlah Kelompok sebanyak 2 Kelompok.....	59
Gambar 4.6 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan <i>Kernel K-Means Clustering</i> dengan jumlah Kelompok sebanyak 3 Kelompok.....	60
Gambar 4.7 Rata-rata Persentase Penduduk dengan Faktor Risiko Penyebab Penyakit Hipertensi menurut Kelompok dengan $k=2$	62
Gambar 4.8 Rata-rata Persentase Penduduk dengan Faktor Risiko Penyebab Penyakit Hipertensi menurut Kelompok dengan $k=3$	64
Gambar 4.9 Rata-rata Persentase Penduduk dengan Faktor Kejiwaan Penyebab Penyakit Hipertensi menurut Kelompok dengan $k=3$	65
Gambar 4.10 Rata-rata Persentase Penduduk dengan Faktor Pola Makan Penyebab Penyakit Hipertensi menurut Kelompok dengan $k=3$	67

Gambar 4.11 Rata-rata Persentase Penduduk dengan Faktor Aktivitas Fisik Penyebab Penyakit Hipertensi menurut Kelompok dengan k=3	69
Gambar 4.12 Rata-rata Persentase Penduduk dengan Faktor Gaya Hidup Penyebab Penyakit Hipertensi menurut Kelompok dengan k=3	70
Gambar 4.13 Rata-rata Persentase Penduduk dengan Faktor Lain-lain Penyebab Penyakit Hipertensi menurut Kelompok dengan k=3	72

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data penelitian	81
Lampiran 2a Hasil Pengelompokan Kabupaten/Kota dengan Metode K-Means dengan k=2.....	82
Lampiran 2b Hasil Pengelompokan Kabupaten/Kota dengan Metode K-Means dengan k=3.....	90
Lampiran 2c Hasil Pengelompokan Kabupaten/Kota dengan Metode Kernel K-Means dengan k=2.....	96
Lampiran 2d Hasil Pengelompokan Kabupaten/Kota dengan Metode Kernel K-Means dengan k=3.....	102
Lampiran 2e Hasil Pengelompokan Kabupaten/Kota menurut Faktor Kejiwaan dengan Metode Kernel K-Means dengan k=3.....	108
Lampiran 2f Hasil Pengelompokan Kabupaten/Kota menurut Pola Makan dengan Metode Kernel K-Means dengan k=3.....	114
Lampiran 2g Hasil Pengelompokan Kabupaten/Kota menurut Aktivitas Fisik dengan Metode Kernel K-Means dengan k=3.....	120
Lampiran 2h Hasil Pengelompokan Kabupaten/Kota menurut Gaya Hidup dengan Metode Kernel K-Means dengan k=3.....	125
Lampiran 2i Hasil Pengelompokan Kabupaten/Kota menurut Obesitas dengan Metode Kernel K-Means dengan k=3.....	131
Lampiran 3 <i>M-files</i> fungsi kernel.....	137
Lampiran 4 <i>Syntax</i> K-Means <i>Clustering</i>	138
Lampiran 5 <i>Syntax</i> Kernel K-Means <i>Clustering</i>	140
Lampiran 6 <i>Syntax</i> Validitas Kelompok untuk Data Simulasi.....	142
Lampiran 7 <i>Syntax</i> Validitas Kelompok untuk Data Penelitian.....	143
Lampiran 8 <i>Syntax</i> Untuk Membangkitkan Data Simulasi.....	144

BAB I

PENDAHULUAN

1.1 Latar Belakang

Salah satu teknik analisis yang dapat menggambarkan karakteristik sekelompok wilayah berdasarkan satu atau lebih variabel antara lain teknik *clustering*. Analisis kelompok adalah analisis statistik yang bertujuan untuk mengelompokkan observasi sedemikian hingga observasi yang berada dalam kelompok yang sama mempunyai sifat yang relatif homogen daripada dalam kelompok yang berbeda (Kaufman, Rousseuw, dan Everitt dalam Hanifah, 2009).

Data clustering banyak diminati berbagai kalangan dan ditulis dalam berbagai paper dan jurnal (Shihab dalam Ramadani, 2010). Maxwell, Pryor dan Smith (2002) dalam Ramadani (2010) pernah menerapkan analisis kelompok dalam penelitian bidang sosial untuk mengelompokkan status sosial masyarakat berdasarkan persamaan dan perbedaan data lintas budaya. Kemudian Wang *et all.* (2005), pernah meneliti tentang penemuan metode baru untuk menggabungkan kelompok dari empat set data *infrared* dari dua tipe kanker menggunakan *teknik clustering* sehingga hasil penelitian mereka menyimpulkan bahwa metode tersebut berpotensi meningkatkan akurasi diagnosa. *Clustering analysis* juga telah digunakan oleh BPS, antara lain pada survey Penyandang Masalah Kesejahteraan Sosial (PMKS), Susenas Modul Perumahan untuk indicator rumah tidak layak huni, dan dalam pengelompokan daerah berdasarkan indicator *Milleium Development Goals* (MDGs).

Analisis pengelompokan sebagai suatu *exploratory procedure* (Johnson dan Wichern: 2002), secara umum dibedakan menjadi metode hirarki dan nonhirarki (metode partisi). *K-means clustering* adalah salah satu metode nonhirarki yang paling banyak digunakan dalam analisis pengelompokan. K-Means merupakan salah satu algoritma pengelompokan konvensional yang umum digunakan karena dinilai efisien digunakan pada data berukuran besar. Metode ini bertujuan membagi n observasi ke dalam K *cluster* dimana tiap observasi menjadi anggota dari *cluster* dengan nilai rata-rata terdekat, melalui proses perulangan (iterasi) hingga pengelompokan tersebut konvergen.

K-Means Clustering merupakan teknik dalam pengelompokan data yang sangat terkenal karena kecepatannya dalam mengelompokkan data. Akan tetapi *K-Means Clustering* memiliki kelemahan didalam memproses data yang berdimensi banyak. Khususnya untuk masukan yang bersifat *non-linearly separable*. *K-Means clustering* juga tidak mampu mengelompokkan data yang bertipe kategorikal dan juga data campuran (numeric dan kategorikal). Kenyataan didunia nyata data yang tersedia atau yang diperoleh memiliki dimensi yang banyak dan juga bersifat campuran. Untuk mengatasi permasalahan ini, telah banyak diusulkan oleh para peneliti metode-metode yang dapat mengatasi kelemahan ini, salah satu diantaranya adalah *Kernel K-Means Clustering* (Dhillon, et. al, 2005).

Kernel K-Means Clustering, pada prinsipnya mirip dengan *K-Means* tradisional, letak perbedaan yang mendasar ada pada perubahan masukannya. Dalam *Kernel K-Means* data *point* akan dipetakan pada dimensi baru yang lebih tinggi menggunakan fungsi non-linier sebelum dilakukan proses *clustering* (Cristianini N, Taylor,J.S dalam Buelens (2009)) . Kemudian *Kernel K-Means* akan mempartisi data menggunakan *linier separator* pada *space* yang baru.

Menurut Santosa (2007), *Kernel K-means clustering* adalah pengembangan dari algoritma *K-means* yang menggunakan metode Kernel untuk memetakan data yang berdimensi tinggi pada *space* yang baru sehingga dapat dipisahkan secara linier. Hal ini dilakukan untuk meningkatkan hasil akurasi pengelompokan. Di dalam *kernel K-means* diharapkan data bisa dipisahkan dengan lebih baik karena data yang *overlap* atau data outlier bisa menjadi linier di ruang dimensi baru. Perluasan dari metode *K-Means* ke *Kernel K-Means* direalisasikan melalui pernyataan jarak dalam bentuk fungsi kernel (Girolami, 2002). Meskipun demikian implementasi tersebut mengalami kendala seperti biaya *clustering* yang tinggi karena kalkulasi yang berulang-ulang dari nilai-nilai kernel, atau memori yang tidak cukup untuk menyimpan matriks kernel. Kendala tersebut berkaitan dengan kompleksitas algoritma *Kernel K-Means*.

Kriteria dalam pengelompokan diantaranya adalah dengan menggunakan ukuran kemiripan, dimana objek dengan kemiripan yang besar akan bergabung membentuk satu kelompok. Ukuran kemiripan yang biasa digunakan antara lain jarak ; *Euclidean*, *Minkowski*, *Squared Euclidean*, *Mahalanobis*, dan *Manhattan*

(Sharma, 1996). Pada umumnya dalam analisis pengelompokan dilakukan dengan menggunakan formula jarak *Euclidean* karena ukuran jarak *euclidean* biasanya digunakan untuk data bertipe numerik. Demikian juga pada paket program MATLAB, jarak yang digunakan adalah *default* jarak Euclidean.

Untuk melihat kinerja metode pengelompokan dengan metode K-Means dan Kernel K-Means pada penelitian ini akan dilakukan studi simulasi dan studi kasus pada data riil. Data simulasi yang digunakan berupa data yang bersifat linier dan non linier. Sedangkan studi kasus pada data riil berupa data persentase penduduk yang memiliki faktor-faktor risiko penyebab penyakit hipertensi.

Hipertensi merupakan suatu keadaan terjadinya peningkatan tekanan darah yang memberi gejala berlanjut pada suatu target organ tubuh sehingga timbul kerusakan lebih berat seperti stroke (terjadi pada otak dan berdampak kematian yang tinggi), penyakit jantung koroner (terjadi kerusakan pembuluh darah jantung). Selain penyakit tersebut, hipertensi dapat pula menyebabkan gagal ginjal, penyakit pembuluh darah lain, diabetes mellitus, dan lain-lain (Sugiharto, 2007).

Di seluruh dunia, hipertensi merupakan masalah yang besar dan serius. Di samping karena prevalensinya yang tinggi dan cenderung meningkat di masa yang akan datang, juga karena tingkat keganasan penyakit yang diakibatkan sangat tinggi seperti penyakit jantung, stroke, gagal ginjal dan lain-lain, juga menimbulkan kecacatan permanen dan kematian mendadak. Kehadiran hipertensi pada kelompok dewasa muda, sangat membebani perekonomian keluarga, karena biaya pengobatan yang mahal dan membutuhkan waktu yang panjang, bahkan seumur hidup (Sugiharto, 2007). Prevalensi penyakit hipertensi di Indonesia semakin memingkat tiap tahunnya. Berdasarkan data Riskesdas, prevalensi penyakit hipertensi di Indonesia sebesar 7,6 persen pada tahun 2007, meningkat menjadi 9,5 persen pada tahun 2013.

Penelitian tentang faktor risiko hipertensi sudah banyak dilakukan sebelumnya, antara lain Kaplan (1985) tentang faktor yang dapat memperbesar risiko atau kecenderungan seseorang menderita hipertensi, Sugiharto (2007) tentang faktor-faktor risiko hipertensi pada masyarakat, studi kasus di Kabupaten Karanganyar menggunakan regresi logistik, Puspita (2009) tentang gaya hidup

mahasiswa penderita hipertensi, Wahyuni (2013) tentang faktor risiko penyakit hipertensi pada laki-laki di wilayah kerja Puskesmas Tawangmangu, Fitiani (2012) tentang pola kebiasaan makan orang lanjut usia, studi kasus penderita penyakit hipertensi suku bangsa Minangkabau di Jakarta.

Penyakit hipertensi sebagai bagian dari penyakit tidak menular (PTM) pada umumnya disebabkan faktor bawaan/keturunan, kecacatan akibat kesalahan proses kelahiran, maupun akibat pola hidup yang tidak sehat, seperti dampak dari konsumsi makanan serta minuman, perilaku merokok, mengonsumsi alkohol, narkoba, kurangnya olah raga, tipe pekerjaan yang banyak duduk, dan pola makanan berkolesterol tinggi serta kurang serat. Faktor-faktor tersebut ditambah lagi dengan perilaku yang serba kompetitif akan meningkatkan stres dan menaikkan tekanan darah. Faktor pencemaran lingkungan seperti asap rokok, asap knalpot, dan asap industri, membuat angka kematian akibat penyakit tidak menular itu meningkat.

Pola hidup yang serba instan membuat masyarakat semakin rentan terkena penyakit tidak menular. Semula PTM seperti hipertensi, stroke, jantung banyak dialami oleh para lansia di atas usia 50 tahun. Namun pola makan dan gaya hidup yang tidak sehat membuat penyakit ini bergeser kepada usia yang lebih muda. Kasus di Samarinda (2010) penyakit stroke diderita remaja berusia 15 tahun, demikian juga kasus di Bali (2014) remaja berusia 15 tahun terserang penyakit jantung koroner.

Masalah PTM pada akhirnya tidak hanya menjadi masalah kesehatan saja, namun bila tidak dikendalikan secara tepat dapat mempengaruhi ketahanan ekonomi nasional karena sifatnya yang kronis dan umumnya mengenai usia produktif. Akibat PTM banyak kerugian yang ditimbulkan. Dari sisi ekonomi, setidaknya terdapat dua kelompok kerugian yang dialami penderita. Pertama adalah kerugian ekonomi yang terbagi menjadi 4 bagian, yaitu dampak penyakit terhadap konsumsi sehat, interaksi sosial, produktivitas jangka pendek dan produktivitas jangka panjang. Kerugian yang kedua adalah adanya dampak penyakit yang mempengaruhi variabel-variabel penting dalam kegiatan ekonomi jangka pendek dan jangka panjang, seperti dampak penyakit terhadap konsumsi,

pendapatan, saving, investasi rumah tangga dan investasi untuk sumber daya manusia (*human capitalinvestment*), (Sugiharto, 2007).

Berdasarkan uraian di atas pada penelitian ini akan dilakukan pengkajian tentang konsep kernel k-means pada data linier dan non linier, serta akan dilakukan pengelompokan kabupaten/kota di Indonesia berdasarkan penduduk dengan faktor risiko penyebab penyakit hipertensi.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan sebelumnya, maka permasalahan yang akan dianalisis dalam tulisan ini adalah bagaimana konsep metode *Kernel K-Means Clustering*. Untuk melihat kinerjanya metode pengelompokan Kernel-K-Means tersebut akan diaplikasikan pada data simulasi yang bersifat linier dan non linier dan pada data riil, yaitu data persentase penduduk dengan faktor risiko penyebab penyakit hipertensi. Selanjutnya hasil pengelompokan dengan metode Kernel K-Means akan dibandingkan dengan hasil pengelompokan menggunakan metode K-Means. Metode pengelompokan yang lebih baik akan digunakan untuk mengelompokkan kabupaten/kota di Indonesia berdasarkan penduduk dengan faktor-faktor risiko penyebab penyakit hipertensi karena untuk mencegah/mengurangi penyakit hipertensi di tiap kabupaten/kota tidak dapat dilakukan dengan kebijakan yang sama di seluruh Indonesia karena karakteristik tiap kabupaten/kota yang berbeda.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang diuraikan di atas, maka tujuan penelitian ini adalah:

1. Mengkaji konsep algoritma *Kernel K-Means*.
2. Membandingkan performansi metode K-Means dan Kernel K-Means pada data simulasi yang bersifat linier dan non linier.
3. Mengelompokkan Kabupaten/Kota di Indonesia berdasarkan penduduk dengan faktor resiko penyebab penyakit hipertensi.

1.4 Manfaat Penelitian

Berdasarkan kajian terhadap kedua teknik pengelompokan di atas, manfaat yang dapat diperoleh dalam penelitian ini adalah :

1. Menambah wawasan keilmuan, khususnya yang berhubungan dengan analisis statistik *multivariate* dalam menganalisis masalah pengelompokan (*clustering*) wilayah.
2. Memberi masukan kepada pemerintah, dalam hal ini terkait dengan program Pengendalian Penyakit Jantung dan Pembuluh Darah (PJPD) untuk melihat wilayah dengan potensi penyakit hipertensi.

1.5 Batasan Masalah

Dalam penelitian ini masalah dibatasi pada faktor risiko yang dapat dikendalikan (*controllable risk factor*) yang mempengaruhi terjadinya penyakit hipertensi pada penduduk usia produktif (15-64 tahun).

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Cluster

Metode statistic untuk pengelompokan observasi dikenal dengan nama analisis kelompok (analisis *cluster*). Analisis kelompok adalah analisis statistic yang bertujuan untuk mengelompokkan observasi sedemikian hingga observasi yang berada dalam kelompok yang sama mempunyai sifat yang relative homogeny daripada dalam kelompok yang berbeda (Kaufman, Rousseuw, dan Everitt dalam Hanifah, 2009).

Analisis *Cluster* merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan kesamaan/kemiripan karakteristik yang dimiliki. Analisis *cluster* mengklasifikasikan objek, dimana setiap objek yang paling dekat kemiripan/kesamaannya dengan objek lain berada dalam suatu *cluster* yang sama. Tujuan dari analisis ini adalah menggabungkan beberapa objek ke dalam kelompok-kelompok (*cluster*) sehingga di dalam setiap kelompok memiliki kemiripan satu sama lain, dan anggota dari suatu kelompok berbeda dengan kelompok lainnya, atau dengan kata lain *cluster-cluster* yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi.

Tahapan dalam melakukan analisis *cluster* menurut Sharma (1996) adalah sebagai berikut :

- a. Menentukan ukuran kemiripan yang digunakan
- b. Menentukan metode pengelompokan, hierarki atau non hierarki
- c. Menentukan teknik yang digunakan untuk metode pengelompokan yang telah ditetapkan
- d. Melakukan analisis *cluster*

2.1.1 Ukuran Kemiripan

Menurut Sharma dalam Dewi (2012), jarak antar objek merupakan faktor yang sangat berpengaruh terhadap hasil kelompok yng dibentuk. Dijelaskan pula

oleh Groenen dalam Reno (2009) bahwa hasil dari kelompok yang akan terbentuk sangat dipengaruhi oleh cara penghitungan jarak antar objek yang digunakan dalam membentuk kelompok. Oleh karena itu perlu ditentukan metode penentuan jarak yang sesuai dengan karakteristik objek yang akan dikelompokkan. Pada paket program MATLAB, jarak yang digunakan adalah default jarak Euclidean.

Berikut adalah beberapa metode pengukuran jarak dalam analisis *cluster* :

1. Jarak Euclidean

Formula jarak Euclidean merupakan tipe jarak yang paling banyak dipilih karena kemudahannya. Dalam penghitungan jarak Euclidean adalah mencari jarak terpendek dari dua titik dengan prinsip orthogonal (tegak lurus). Formula jarak Euclidean antara dua objek (x_i dan x_j), yang berdimensi p adalah sebagai berikut :

$$d_{ij}(x_i, x_j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{pi} - x_{pj})^2} \quad (2.1)$$

Dimana :

x_i : Vektor pengamatan ke-i

x_j : Vektor pengamatan ke-j

Formula jarak Euclidean ini memiliki keuntungan, yaitu jarak dari dua objek apa saja tidak dipengaruhi oleh penambahan objek baru untuk dianalisis, yang mungkin merupakan pencilan. Jarak Euclidean biasanya dihitung dari data mentah, bukan data yang telah distandarkan, sehingga memiliki kelemahan yaitu pada saat observasi yang digunakan memiliki skala pengukuran yang berbeda tidak tepat jika menggunakan jarak Euclidean. Alternative yang biasa ditempuh adalah menggunakan formula jarak *Standardized Euclidean* atau Mahalanobis (Sharma, 1996).

2. Jarak Mahalanobis

Formula jarak Mahalanobis digunakan saat data yang dimiliki memiliki perbedaan dalam skala pengukuran, selain itu terdapat korelasi pada data. Formula jarak Mahalanobis dinyatakan sebagai berikut :

$$d_{ij}^2(x_i, x_j) = (x_i - x_j) \mathbf{V}^{-1} (x_i - x_j)^T \quad (2.2)$$

Dimana :

x_i : Vektor pengamatan ke-i

x_j : Vektor pengamatan ke-j

V : Matriks covarians dari x

3. Jarak Minkowski

Minkowski merupakan jarak yang didasarkan pada akar dari jumlah perbedaan nilai absolut. Formula jarak Minkowski adalah sebagai berikut :

$$d_{ij}^2(x_i, x_j) = \left(\sum_{k=1}^n |x_{ni} - x_{nj}|^p \right)^{\frac{1}{p}} \quad (2.3)$$

Dengan :

x_i : Vektor pengamatan ke-i

x_j : Vektor pengamatan ke-j

P : Parameter dari formula jarak Minkowski

Keuntungan dari pengembangan algoritma menggunakan jarak Minkowski adalah penurunan matematisnya dapat digunakan pada berbagai formula jarak, dengan memodifikasi parameter p yang terdapat dalam formula jarak Minkowski (Groenen, et al dalam Munaf, 2010). Misalnya pada saat $p=1$, formula jarak Minkowski sama dengan formula jarak Manhattan atau *city-block*, dan pada saat $p=2$, sama dengan formula jarak Euclidean.

2.1.2 Metode Pengelompokan

Setelah mendapatkan ukuran kemiripan, langkah selanjutnya adalah menentukan metode pengelompokan. Metode pengelompokan yang lazim digunakan dalam analisis kelompok (cluster) adalah metode hierarki dan metode nonhierarki.

Metode pengelompokan hirarki digunakan untuk mengelompokkan objek secara terstruktur berdasarkan ukuran kemiripan, dimana objek yang mempunyai jarak terdekat kemudian digabungkan menjadi satu kelompok, dan kelompok yang diinginkan belum diketahui banyaknya. Metode pengelompokan hirarki dapat dilakukan dengan dua pendekatan, yaitu dengan cara penggabungan

(*agglomerative*), dan pemisahan (*divisive*). Metode hirarki dengan pendekatan *agglomerative* didapat dengan menggabungkan objek secara bertahap, dimana pada awalnya tiap-tiap objek dianggap sebagai kelompok tersendiri, kemudian kelompok-kelompok dengan kemiripan terbesar bergabung yang pada akhirnya didapat satu kelompok saja. Ada lima metode *agglomerative* dalam pembentukan kelompok, yaitu : pautan tunggal (*single linkage*), pautan lengkap (*complete linkage*), pautan rata-rata (*average linkage*), metode ward (*ward method*), dan metode centroid (*centroids method*). Sebaliknya metode hirarki dengan pendekatan *divisive* dimulai dengan membentuk satu kelompok besar beranggotakan seluruh objek. Kelompok besar tersebut kemudian dipisah menjadi kelompok yang lebih kecil, sampai satu kelompok hanya beranggotakan satu objek saja. Kedua cara dalam metode hirarki ini tidak berbeda dalam pembentukan kelompok yang terjadi, tetapi hanya berbeda dalam tahapan pembentukan kelompok saja.

Metode pengelompokan nonhirarki digunakan jika banyaknya kelompok yang akan dibentuk sudah diketahui. Ada dua hal yang harus diketahui dalam pengelompokan nonhirarki, yaitu jumlah kelompok dan pemilihan pusat kelompok (*centroid*). Proses pengelompokan nonhirarki dimulai dengan menentukan jumlah kelompok yang akan dibentuk, kemudian menghitung jarak masing-masing objek dengan pusat kelompok (*centroid*). Kemudian menghitung rata-rata *centroid* untuk kelompok yang dibentuk. Proses ini dilakukan terus menerus sampai tidak terjadi lagi pemindahan objek. Pengelompokan dengan metode nonhirarki lebih cepat daripada metode hirarki, dan lebih menguntungkan jika jumlah objek atau observasinya besar (sampel besar). Namun lebih lanjut, Sharma menjelaskan bahwa pengelompokan dengan metode hirarki dan nonhirarki sebaiknya digunakan untuk saling melengkapi, bukan untuk dibandingkan.

2.1.3 Metode K-Means Clustering

Metode K-means merupakan salah satu metode pengelompokan data (*clustering*) nonhirarki. Algoritma K-means lebih sering digunakan karena lebih sederhana dan cepat prosesnya. MacQueen dalam Johnson (2007) menyarankan

penggunaan K-means untuk menjelaskan algoritma dalam penentuan suatu objek ke dalam suatu kelompok tertentu berdasarkan rata-rata terdekat.

Pada K-means, misal didefinisikan $X = \{x_1, \dots, x_n\}$ adalah sebuah himpunan data dalam ruang berdimensi D , yang dinotasikan R^D , sedangkan k adalah sebuah bilangan integer positif lebih dari satu. $x_n \in R^D$, maka algoritma K-means *clustering* akan membagi (mempartisi) X ke dalam k kelompok, dapat dinyatakan dengan himpunan X_1, X_2, \dots, X_k yang saling lepas, sehingga $X_1 \cup X_2 \cup \dots \cup X_k = X$, dimana setiap kelompok memiliki nilai tengah (*centroid*) dari data-data dalam kelompok tersebut. Algoritma K-means secara acak menentukan k buah data sebagai titik tengah (*centroid*), kemudian dengan pendekatan jarak Euclidean dihitung jarak antara data dengan *centroid*, untuk selanjutnya data akan ditempatkan ke dalam kelompok yang terdekat dihitung dari titik tengah kelompok. Proses penentuan *centroid* dan penempatan data kelompok dilakukan sampai nilai *centroid*-nya konvergen (*centroid* dari semua kelompok tidak berubah lagi).

Dalam bentuk yang paling sederhana langkah-langkah algoritma k-means terdiri dari tiga tahap :

1. Bagi objek-objek ke dalam K kelompok dan tentukan pula *centroid*-nya di tiap kelompok
2. Masukkan objek ke suatu kelompok berdasarkan nilai rata-rata terdekat. Jarak yang digunakan biasanya menggunakan jarak Euclidean, dan hitung kembali rata-rata untuk kelompok baru yang terbentuk
3. Ulangi langkah 2 sampai tidak ada lagi pemindahan objek antar kelompok.

Cara yang sederhana untuk menentukan kualitas kelompok adalah dengan mengukur seberapa dekat masing-masing titik-titik kelompok dengan *centroid*, dengan demikian ukuran kepadatan kelompok dinyatakan sebagai total jarak dari rata-rata kelompok.

$$\sum_{x_n \in C_k} \|x_n - m_k\|^2 = \sum_{n=1}^N z_{kn} \|x_n - m_k\|^2 \quad (2.4)$$

Dimana pusat *cluster* didefinisikan sebagai :

$$\mathbf{m}_k = \frac{1}{N_k} \sum_{x_n \in C_k} x_n \quad (2.5)$$

Dengan z_{kn} menunjukkan kenggotaan data x_n terhadap kelompok k, dimana $z_{kn} = 1$ jika $x_n \in C_k$, dan 0 untuk yang lainnya dan $N_k = \sum_{n=1}^N z_{kn}$ adalah total banyaknya titik yang dialokasikan ke kelompok k. Maka kebaikan kelompok berdasarkan jumlah ukuran kepadatan kelompok untuk masing-masing kelompok k. dengan menggunakan indicator variable z_{kn} bisa ditentukan ukuran kebaikan kelompok secara keseluruhan sebagai berikut :

$$\mathcal{E}_K = \sum_{n=1}^N \sum_{k=1}^K z_{kn} \|x_n - \mathbf{m}_k\|^2 \quad (2.6)$$

Setelah memperoleh ukuran secara keseluruhan dari kualitas kelompok, langkah selanjutnya adalah menyusun algoritma yang akan memungkinkan untuk mengoptimalkan ukuran ini. Kriteria yang harus dioptimalkan adalah $\mathcal{E}_K = \sum_{n=1}^N \sum_{k=1}^K z_{kn} \|x_n - \mathbf{m}_k\|^2$, dimana terdapat dua set parameter, yaitu rata-rata kelompok (\mathbf{m}_k) dan variable indikator alokasi cluster z_{kn} . Optimalisasi pada setiap set variabel dengan memegang set yang bersifat tetap (*fixed*) sementara melakukan optimasi pada set lain.

Kriteria pengklusteran ketika menggunakan algoritma K-Mean adalah meminimalkan nilai *clustering error* (\mathcal{E}_K), atau secara sederhana sebagai berikut :

$$\mathcal{E}_K = \sum_{n=1}^N \sum_{k=1}^K z_{kn} \|X_n - \mathbf{m}_k\|^2$$

Dapat dijabarkan sebagai berikut :

$$\begin{aligned} \mathcal{E}_K &= \sum_{n=1}^N \sum_{k=1}^K z_{kn} \|X_n - \mathbf{m}_k\|^2 \\ &= \sum_{n=1}^N \sum_{k=1}^K z_{kn} (X_n - \mathbf{m}_k)^T (X_n - \mathbf{m}_k) \\ &= \sum_{n=1}^N \sum_{k=1}^K z_{kn} (X_n^T X_n - 2\mathbf{m}_k^T X_n + \mathbf{m}_k^T \mathbf{m}_k) \end{aligned} \quad (2.7)$$

Dengan catatan :

$$\mathbf{m}_k^T \mathbf{X}_n = \frac{1}{N_k} \sum_{m=1}^N z_{km} \mathbf{X}_m^T \mathbf{X}_n \quad (2.8)$$

Dan

$$\begin{aligned} \mathbf{m}_k^T \mathbf{m}_k &= \left(\frac{1}{N_k} \sum_{p=1}^N Z_{kp} \mathbf{X}_p \right)^2 \\ &= \frac{1}{N_k^2} \sum_{p=1}^N \sum_{l=1}^N Z_{kp} Z_{kl} \mathbf{X}_p^T \mathbf{X}_l \end{aligned} \quad (2.9)$$

2.1.4 Metode Kernel K-Means Clustering

Fungsi Kernel

Fungsi kernel adalah suatu fungsi K dimana untuk semua vektor input \mathbf{x}, \mathbf{z} akan memenuhi kondisi :

$$K(\mathbf{x}, \mathbf{z}) = \phi(\mathbf{x})^T \phi(\mathbf{z}) \quad (2.10)$$

Dimana $\phi(\cdot)$ adalah fungsi pemetaan dari ruang input (*input space*) ke ruang fitur (*feature space* F), atau secara matematis $\phi: \mathbf{x} \rightarrow \phi(\mathbf{x}) \in F$. Fungsi kernel memungkinkan untuk mengimplementasikan suatu model pada ruang berdimensi lebih tinggi (ruang fitur) tanpa harus mendefinisikan fungsi pemetaan dari ruang input ke ruang fitur, sehingga untuk kasus yang *nonlinearly separable* pada ruang input diharapkan akan menjadi *linearly separable* pada ruang fitur dengan menggunakan hyperplane sebagai *decision boundary* secara efisien.

Menurut Cristiani dkk (2004) dalam Murfi (2009), fungsi kernel ada 4 macam yaitu :

- a. Kernel Linier

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$$

- b. Polynimial Kernel

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d, \gamma > 0$$

- c. Fungsi Kernel Gaussian (*Radial Basis Function*)

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp \left\{ -\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2 \right\}$$

d. Eksponensial (sigmoid) Kernel

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r)$$

Dengan γ, r , dan d merupakan parameter kernel, dan $i, j = 1, 2, \dots, n$. Fungsi kernel yang biasa digunakan adalah kernel Gaussian karena kernel Gaussian lebih mudah dalam perhitungan dan lebih sederhana dalam komputasinya.

2.1.4.2 Sifat-sifat Kernel

Fungsi kernel secara umum memiliki sifat sebagai berikut :

a. Matriks kernel adalah berupa matrik simetris, dimana

$$K(\mathbf{x}_i, \mathbf{x}_j) = K(\mathbf{x}_j, \mathbf{x}_i)$$

b. Matriks $N \times N$ dengan elemen matriks $K(\mathbf{x}_i, \mathbf{x}_j) \forall i, j = 1, \dots, N$ merupakan matriks *positif semi definite*.

Sedangkan sifat lain dari Kernel untuk sembarang ruang sampel X dengan kernel k_1 dan k_2 yang melalui X , antara lain adalah :

1. $k(x, y) = k_1(x, y) + k_2(x, y)$
2. $k(x, y) = ak_1(x, y)$, dimana $a > 0$
3. $k(x, y) = f(x) \cdot f(y)$ untuk sembarang fungsi f terhadap x
4. $k(x, y) = k_1(x, y) \cdot k_2(x, y)$
5. $k(x, y) = \frac{k_1(x, y)}{\sqrt{k_1(x, x)} \sqrt{k_2(y, y)}}$

Kernel K-Means Clustering

Menurut Santosa (2007), *Kernel K-means clustering* adalah pengembangan dari algoritma *K-means* yang menggunakan metode Kernel untuk memetakan data yang berdimensi tinggi pada *space* yang baru sehingga dapat dipisahkan secara linier.

Pada pengelompokan dengan *Kernel K Means*, masing-masing element dari ε_k dapat dituliskan dengan istilah sebagai *inner product* dan dengan

menggunakan fungsi pemetaan $\phi(\cdot)$ data dipetakan dari *input space* ke dalam *featur space*, kemudian dengan mengambil istilah diatas dan memasukkannya mereka secara bersamaan kita dapat menuliskan kriteria kluster dalam *featur space* yang didefinisikan oleh ϕ sebagai ε_k^ϕ , di mana vektor *mean* yang telah dipetakan dalam *space* ini dinyatakan sebagai $\phi(m_k)$. Sehingga persamaannya menjadi seperti di bawah ini :

$$\varepsilon_K^\phi = \sum_{n=1}^N \sum_{k=1}^K Z_{kn} \|\phi(X_n) - \phi(m_k)\|^2 \quad (2.11)$$

Dimana :

$$\phi(m_k)^T \phi(X_n) = \frac{1}{N_k} \sum_{m=1}^N z_{km} \phi(X_m)^T \phi(X_n) \quad (2.12)$$

dan

$$\begin{aligned} \phi(m_k)^T \phi(m_k) &= \left(\frac{1}{N_k} \sum_{p=1}^N z_{kp} \phi(X_p) \right)^2 \\ &= \frac{1}{N_k^2} \sum_{p=1}^N \sum_{l=1}^N z_{kp} z_{kl} \phi(X_p)^T \phi(X_l) \end{aligned} \quad (2.13)$$

Ketika jarak antara data ke n dengan pusat *cluster* ke- k dalam ruang berdimensi tinggi dilambangkan dengan δ_{kn} , maka:

$$\delta_{kn} = \|\phi(X_n) - \phi(m_k)\|^2$$

$$\text{Sehingga } \varepsilon_K^\phi = \sum_{n=1}^N \sum_{k=1}^K z_{kn} \delta_{kn}$$

Point pertama yang harus diperhatikan adalah kriteria *cluster* dapat dituliskan hanya dari fungsi kernel yang dihitung pada masing-masing pasangan titik data, dengan demikian suatu algoritma dapat dikembangkan dimana hanya membutuhkan suatu tahapan dalam memperbarui variabel indikator z_{kn} karena tidak ada pembaharuan secara *explicit* dari nilai *mean cluster* yang digunakan.

2.1.5 Evaluasi Hasil Pengelompokan

Pendekatan yang umum dipakai untuk mengevaluasi kebaikan dari hasil pengelompokan dikenal dengan *cluster validation* (Maulik, et al, 2002). *Cluster*

Validation yang akan dibahas disini dibatasi pada validitas kelompok dengan pendekatan *internal clustering validation*, yaitu evaluasi hasil pengelompokan tanpa informasi dari luar, dalam hal ini berdasarkan seberapa dekat jarak di dalam kelompok dan jarak antar kelompok.

Liu, et al (2010) melakukan pengujian kinerja 11 ukuran validasi terhadap beberapa aspek yang berbeda, yaitu terhadap dampak kemonotonan (*monotonicity*), dampak *noise*, dampak kepadatan (*density*), dampak adanya *subcluster*, dan dampak terhadap kemencengan distribusi (*skewed distribution*). Jenis-jenis *internal clustering validation* yang digunakan pada penelitian Liu, et al (2010) adalah :

1. *R Squared (RS)*

RS mengukur ketidakseragaman kelompok (*dissimilarity*). Nilai RS berada pada interval 0-1, dimana bernilai 0 jika tidak terdapat perbedaan antar kelompok, dan bernilai 1 jika kelompok berbeda secara signifikan. Secara matematis diuraikan berikut ini :

$$RS = \frac{SS_l - SS_w}{SS_l}, \quad (2.14)$$

Dimana :

$$SS_l = \sum_{j=1}^d \sum_{k=1}^{n_j} (x_k - \bar{x}_j)^2, SS_w = \sum_{j=1, \dots, d}^{i=1, \dots, n_c} (x_k - \bar{x}_j)^2$$

n_c : banyaknya kelompok

d : dimensi

2. *Root Mean Square Standar Deviation (RMSSTD) Index*

Menurut Sharma (1996), RMSSTD adalah varians dari kelompok, yang mengukur kehomogenan dari kelompok. Sebagaimana tujuan pengelompokan yaitu mengidentifikasi keseragaman dalam kelompok, maka semakin kecil nilai RMSSTD semakin baik.

$$RMSSTD = \sqrt{\frac{\sum_{j=1, \dots, d}^{i=1, \dots, n_c} \sum_{k=1}^{n_j} (x_k - \bar{x}_j)^2}{\sum_{j=1, \dots, d}^{i=1, \dots, n_c} (n_{ij} - 1)}} \quad (2.15)$$

3. *Modified Hubert Γ Statistic*

Indeks ini menggambarkan derajat pemisahan dari data set, dengan catatan hanya ketika dua data point berada pada kelompok yang berbeda akan dipengaruhi oleh besarnya nilai Gamma (Γ), jika tidak maka tidak berpengaruh pada nilai Gamma (Γ). Secara matematis indeks *Modified Hubert Γ Statistic* adalah :

$$\Gamma = \frac{1}{m} \sum_{i=1}^{n-1} \sum_{j=i+1}^n P(i, j) Q(i, j) \quad (2.16)$$

Dimana n adalah banyaknya objek di dalam data set, $m = n(n-1)/2$, P adalah matriks data set, dengan elemen (i, j) , jarak antara titik data x_i dan x_j , dan Q adalah matriks ukuran $n \times n$ dimana elemen (i, j) sama dengan jarak antara pusat kelompok dengan masing-masing data x_i dan x_j .

4. *Calinski-Harabasz Index*

Indeks ini mengukur jumlah kuadrat antar kelompok dan jumlah kuadrat di dalam kelompok. Untuk menentukan banyaknya jumlah kelompok (n_c) dengan menggunakan nilai maksimum dari nilai Calinski-Harabasz CH_{n_c} .

$$CH_{n_c} = \frac{SS_l / (n_c - 1)}{SS_w / (n - n_c)}$$

Dengan n_c banyaknya kelompok, SS_l jumlah kuadrat jarak antar kelompok, SS_w jumlah kuadrat jarak di dalam kelompok,

5. *I Index*

$$I = \frac{1}{NC} \frac{\sum_{x \in D}^i d(x, c)}{\sum_i \sum_{x \in C_i} d(x, c_i)} \max_{i,j} d(c_i, c_j)^p$$

Dengan NC adalah jumlah kelompok; C_i kelompok ke- i , dan c_i adalah pusat dari kelompok ke C_i . I indeks akan memperoleh nilai maksimum ketika sudah diperoleh banyaknya kelompok yang sesuai.

6. *Dunn's Indices*

Indeks Dunn digunakan untuk menentukan pemisah yang baik dan kelompok yang baik, yang merupakan rasio nilai jarak minimum di dalam kelompok (*intra cluster*) dan jarak maksimum antar kelompok (*inter cluster*).

$$Dunn_{index} = \frac{Dist_{min}}{Dist_{max}}$$

Dimana $Dist_{min}$ adalah jarak minimum di dalam kelompok, dan $Dist_{max}$ adalah jarak maksimum antar kelompok. Nilai maksimum indeks Dunn menunjukkan nilai kelompok yang optimal.

7. *Silhouette Index*

S_{index} untuk menghitung rata-rata ketidakmiripan antar titik dalam kelompok yang sama dan kelompok yang berbeda.

$$S_{index} = \frac{b_{(i)} - a_{(i)}}{\max\{a_{(i)}, b_{(i)}\}}$$

Dimana :

$a_{(i)}$: rata-rata jarak objek ke- i terhadap semua objek dalam kelompok yang sama

8. *Davies-Bouldin Index*

DB_{index} digunakan untuk menghitung rasio jumlah sebaran di dalam kelompok dan antar kelompok. Nilai DB_{index} yang kecil menunjukkan jumlah kelompok yang terbaik.

$$DB_{index} = \frac{1}{k} \sum_{i=1, i \neq j}^k \max \left(\frac{averdis_i + averdis_j}{dis(c_i, c_j)} \right)$$

Dimana :

k : banyaknya kelompok

$averdis_i$: rata-rata jarak semua objek di dalam kelompok (cluster) i terhadap pusat kelompok (cluster) c_i

$averdis_j$: rata-rata jarak semua objek di dalam kelompok (cluster) j terhadap pusat kelompok (cluster) c_j

$dis(c_i, c_j)$: jarak antara pusat kelompok c_i dan c_j

9. Xie-Beni Index

Indeks *Xie-Beni* digunakan untuk mengukur validitas kelompok pada *Fuzzy Clustering* (Xie-Beni, 1991).

10. SD Validity Index

Indeks validitas SD berdasarkan rata-rata sebarankelompok, dan total pemisahan kelompok. Rata-rata sebaran kelompok didefinisikan sebagai :

$$Scatt(n_c) = \frac{1}{n_c} \sum_{i=1, \dots, n_c} \frac{\|\sigma(v_i)\|}{\|\sigma(X)\|} \quad (2.17)$$

Sedangkan total pemisahan antar kelompok didefinisikan sebagai :

$$Dis(n_c) = \frac{D_{\max}}{D_{\min}} \sum_{k=1, \dots, n_c} \left[\sum_{z=1, \dots, n_c} \|v_k - v_z\| \right]^{-1} \quad (2.18)$$

Dimana $D_{\max} = \max(\|v_i - v_j\|), \forall i, j \in \{1, 2, 3, \dots, n_c\}$ adalah jarak maksimum

antar pusat kelompok, dan $D_{\min} = \min(\|v_i - v_j\|), \forall i, j \in \{1, 2, 3, \dots, n_c\}$ jarak

minimum antar pusat kelompok. Sehingga indeks validitas SD adalah :

$$(SD(n))_{index} = a.Scatt(n_c) + Dis(n_c) \quad (2.19)$$

Dengan a adalah faktor penimbang, sama dengan $Dis(c_{\max})$, dengan c_{\max} adalah maksimum jumlah input kelompok.

11. S_Dbw Validity Index

Indeks validitas *S_Dbw* prinsipnya sama dengan indeks SD, yaitu berdasarkan kekompakan kelompok, tetapi juga memperhatikan kepadatan kelompok. *S_Dbw* mengukur varians di dalam kelompok (*intra cluster variance*) dan varian antar kelompok (*inter cluster variance*). Varians di dalam kelompok mengukur rata-rata sebaran kelompok, sebagaimana dijabarkan pada rumus berikut :

$$Scatt = \frac{1}{n} \sum_{i=1}^{n_c} \frac{\|\sigma(v_i)\|}{\|\sigma(x)\|} \quad (2.20)$$

Sedangkan Kepadatan di dalam kelompok didefinisikan sebagai berikut :

$$Dens_bw = \frac{1}{n_c(n_c - 1)} \sum_{i=1}^{n_c} \left(\sum_{\substack{j=1 \\ i \neq j}}^{n_c} \frac{density(u_{ij})}{\max\{density(v_i), density(v_j)\}} \right) \quad (2.21)$$

Dimana u_{ij} adalah titik tengah garis pemisah yang didefinisikan oleh v_i dan v_j pusat kelompok. Fungsi kepadatan di sekitar titik didefinisikan sebagai banyaknya titik dalam kelompok dengan jarak yang sama terhadap standar deviasi kelompok. Rata-rata standar deviasi kelompok didefinisikan sebagai berikut :

$$stdev = \frac{1}{n_c} \sqrt{\sum_{i=1}^{n_c} \|\sigma(v_i)\|} \quad (2.22)$$

Sehingga S_Dbw indeks adalah :

$$S_Dbw = Scatt + Dens_bw \quad (2.23)$$

Berikut adalah hasil penelitian Liu, dkk (2010) tentang perbandingan kinerja 11 ukuran indeks validitas :

Tabel 2.1 Evaluasi Performansi 11 Ukuran Validitas

Ukuran Validitas	Evaluasi terhadap dampak				
	Monoton	Noise	Kepadatan	Sub Cluster	Kemencengan
RMSSTD	√	-	-	-	-
RS	√	-	-	-	-
Γ	√	-	-	-	-
CH	-	√	-	-	√
I	-	-	√	-	-
D	-	√	-	√	-
S	-	-	-	√	-
DB**	-	-	-	√	-
SD	-	-	-	√	-
S_Dbw	-	-	-	-	-
XB**	-	-	-	√	-

Berdasarkan tabel evaluasi di atas, dari 11 performa *internal clustering validation measure* hanya satu ukuran yang mampu mengatasi seluruh kendala dalam pengelompokan dengan K-means Clustering, yaitu metode *S_Dbw* index.

Dalam penelitian ini, metode *internal clustering validation* yang digunakan menggunakan indeks *S_Dbw*.

2.2 Faktor Risiko Penyakit Hipertensi

Faktor resiko adalah semua faktor penyebab ditambah dengan faktor biologis yang berhubungan dengan penyakit. Faktor resiko merupakan faktor-faktor yang ada sebelum terjadinya penyakit (Bustan dalam Nababan, 2008). Faktor resiko adalah bagian dari ilmu epidemiologi. Epidemiologi adalah ilmu yang mempelajari pola kesehatan dan penyakit, serta faktor yang terkait di tingkat populasi. Epidemiologi pada penyakit menular disebut etiologi, sedangkan pada penyakit tidak menular disebut sebagai faktor resiko.

Hipertensi merupakan suatu penyakit kronis yang sering disebut *silent killer* karena pada umumnya pasien tidak mengetahui bahwa mereka menderita penyakit hipertensi sebelum memeriksa tekanan darahnya. Selain itu penderita hipertensi umumnya tidak mengalami suatu tanda atau gejala sebelum terjadi komplikasi.

Pengertian hipertensi menurut WHO adalah keadaan seseorang apabila mempunyai tekanan sistolik sama dengan atau lebih tinggi dari 160 mmHg dan tekanan diastolic sama dengan atau lebih tinggi dari 80 mmHg secara konsisten dalam beberapa waktu. Batasan ini tidak membedakan antara usia dan jenis kelamin (Marliani dkk, 2007). Hipertensi adalah suatu gangguan pada pembuluh darah yang mengakibatkan suplai oksigen dan nutrisi yang dibawa darah terhambat sampai ke jaringan tubuh yang membutuhkannya (Sustrani, 2006). Hipertensi merupakan salah satu faktor resiko terbesar penyebab morbiditas dan mortalitas pada penyakit kardiovaskular (Kearney et al., 2005)

Terdapat faktor yang mempengaruhi jumlah hipertensi seperti ras, umur, obesitas, asupan garam yang tinggi, dan adanya riwayat hipertensi dalam keluarga. Menurut Kaplan (1985) faktor yang dapat memperbesar risiko atau kecenderungan seseorang menderita hipertensi, diantaranya ciri-ciri individu

seperti umur, jenis kelamin dan suku, faktor genetik serta faktor lingkungan yang meliputi obesitas, stres, konsumsi garam, merokok, konsumsi alkohol, dan sebagainya. Selama tahun 2010, 1,7 juta kematian pada penyakit kardiovaskuler disebabkan karena konsumsi garam yang berlebihan (WHO, 2015). Menurut Tabrani (1995) dalam Puspita WR (2009) komplikasi hipertensi antara lain: penyakit jantung, stroke, gagal ginjal, kelainan mata, dan diabetes mellitus.

Menurut Sugiharto (2007), faktor risiko penyakit hipertensi dibedakan atas :

a. Faktor risiko yang tidak dapat diubah/dikendalikan

1. Umur

Hipertensi erat kaitannya dengan umur, semakin tua seseorang semakin besar risiko terserang hipertensi. Umur lebih dari 40 tahun mempunyai risiko terkena hipertensi. Dengan bertambahnya umur risiko terkena hipertensi lebih besar, sehingga prevalensi hipertensi di kalangan usia lanjut cukup tinggi

2. Jenis Kelamin

Bila ditinjau dari perbandingan antara wanita dan pria, terdapat angka yang cukup bervariasi. Menurut Mansjoer dalam Sugiharto (2007), pria dan wanita yang telah menopause mempunyai peluang yang sama untuk terjadi hipertensi. Namun menurut Bustan dalam Sugiharto, wanita lebih banyak yang menderita hipertensi dibandingkan pria, dalam hal ini terkait dengan terdapatnya hormon estrogen pada wanita.

3. Riwayat Keluarga

Seseorang dengan sejarah keluarga yang mempunyai riwayat hipertensi lebih sering menderita hipertensi. Keluarga yang memiliki hipertensi dan penyakit jantung meningkatkan risiko hipertensi 2-5 kali lipat.

4. Genetik

Peran faktor genetik terhadap timbulnya hipertensi terbukti dengan ditemukannya kejadian bahwa hipertensi lebih banyak pada kembar monozigot (satu sel telur) daripada heterozigot (berbeda sel telur).

b. Faktor risiko yang dapat diubah/dikendalikan

1. Kebiasaan merokok

Hubungan antara rokok dengan peningkatan risiko penyakit kardiovaskuler telah banyak dibuktikan. Selain lamanya, banyaknya konsumsi rokok juga meningkatkan risiko hipertensi.

2. Konsumsi asin/garam

Garam merupakan hal yang sangat penting pada mekanisme timbulnya hipertensi. Hipertensi hampir tidak pernah ditemukan pada suku bangsa dengan asupan garam minimal.

3. Konsumsi lemak jenuh

Kebiasaan konsumsi lemak jenuh erat kaitanya dengan peningkatan berat badan yang berisiko terjadinya hipertensi. Konsumsi lemak jenuh juga meningkatkan risiko aterosklerosis yang berkaitan dengan kenaikan tekanan darah.

4. Penggunaan jelantah

Jelantah adalah minyak goreng yang sudah lebih dari satu kali pakai untuk menggoreng, dan minyak goreng ini merupakan minyak yang telah rusak.

5. Kebiasaan konsumsi minum-minuman beralkohol

Peminum alkohol berat cenderung terkena hipertensi dibandingkan seseorang yang tidak minum atau minum sedikit.

6. Obesitas

Obesitas adalah kondisi kegemukan, dimana berat badan lebih besar dari indeks massa tubuh yang ideal.

7. Olahraga (aktifitas fisik)

Kurangnya aktifitas fisik meningkatkan risiko menderita hipertensi karena meningkatkan risiko kelebihan berat badan. Orang yang tidak aktif juga cenderung mempunyai denyut jantung yang lebih tinggi sehingga otot jantungnya harus bekerja lebih keras pada setiap kontraksi.

8. Stress

Stres atau ketegangan jiwa (rasa tertekan, murung, bingung, cemas, berdebar-debar, rasa marah, dendam, rasa takut, rasa bersalah) dapat

merangsang kelenjar melepaskan hormon adrenalin yang dapat memicu jantung berdenyut lebih cepat serta lebih kuat, sehingga tekanan akan meningkat.

9. Penggunaan estrogen

Menurut Bustan dalam Sugiharto (2007), lamanya pemakaian kontrasepsi estrogen (kurang lebih 12 tahun berturut-turut) akan meningkatkan tekanan darah perempuan.

BAB III

METODOLOGI PENELITIAN

Metodologi penelitian dalam penelitian ini terdiri dari kajian teori dan kajian terapan. Adapun uraian selengkapnya adalah sebagai berikut :

3.1 Kajian Teori

Bagian ini akan mengkaji konsep kernel k-means *clustering* sebagaimana berikut :

1. Mengkaji sifat-sifat kernel

Sifat-sifat kernel yang akan dikaji adalah :

- a. Matrik kernel merupakan matriks simetris
- b. Matriks kernel merupakan matriks *positif semi definite*
 - a. $k(x, y) = k_1(x, y) + k_2(x, y)$
 - b. $k(x, y) = ak_1(x, y)$, dimana $a > 0$
 - c. $k(x, y) = f(x).f(y)$ untuk sembarang fungsi f terhadap x
 - d. $k(x, y) = k_1(x, y) \cdot k_2(x, y)$
 - e. $k(x, y) = \frac{k_1(x, y)}{\sqrt{k_1(x, x)}\sqrt{k_2(y, y)}}$

2. Mengkaji konsep algoritma kernel k-means

3. Mengkaji kompleksitas algoritma kernel k-means

3.2 Kajian Terapan

Data yang digunakan dalam kajian terapan ini terdiri dari dua tipe data, yaitu data simulasi dan data riil yang berupa data sekunder, yaitu data simulasi dari penelitian Epa Suryanto (2015), dan data Riset Kesehatan Dasar (RISKESDAS) tahun 2013 yang bersumber dari Badan Penelitian dan Pengembangan Kesehatan (Balitbangkes) Kementerian Kesehatan RI. *Software* yang digunakan adalah SPSS, Matlab, dan R.

3.2.1 Penerapan pada Data Simulasi

Data simulasi yang digunakan terdiri dari 2 jenis data, yaitu data yang bersifat linier dan non linier. Variabel yang digunakan terdiri dari dua variabel, yaitu variabel x_1 dan x_2 yang selanjutnya akan dikelompokkan ke dalam 2 kelompok yang berbeda. Desain data simulasi yang digunakan dalam penelitian ini adalah :

a. Data linier

- Menentukan banyaknya observasi (n), sebanyak 500, 1000, dan 3000.
- Data dibangkitkan dari distribusi normal dengan rata-rata dan varians yang sama, yaitu rata-rata 0 dan varians 1.
- Variabel x_1 dan x_2 dikelompokkan ke dalam dua kelas yang berbeda
- Berdasarkan program untuk membangkitkan data simulasi sebagaimana tersaji pada Lampiran.8a, fungsi untuk mengelompokkan data simulasi berupa data linier adalah :

$$f(x_1, x_2) = \begin{cases} 1, & x_1 > x_2 \\ 2, & x_1 \leq x_2 \end{cases}$$

- apabila nilai x_1 lebih besar daripada nilai x_2 maka data tersebut berada di kelompok 1, dan jika nilai x_1 kurang dari atau sama dengan nilai x_2 maka data tersebut berada di kelompok 2.

b. Data non linier (lingkaran)

- Menentukan banyaknya observasi (n), sebanyak 500, 1000, dan 3000.
- Data dibangkitkan dari distribusi normal dengan rata-rata dan varians yang sama, yaitu rata-rata 0 dan varians 1.
- Variabel x_1 dan x_2 dikelompokkan ke dalam dua kelas yang berbeda
- Berdasarkan program untuk membangkitkan data simulasi sebagaimana tersaji pada Lampiran.8b, fungsi untuk mengelompokkan data simulasi berupa data non linier adalah :

$$f(x_1, x_2) = \begin{cases} 1, & \sqrt{x_1^2 + x_2^2} < 1.2 \\ 2, & \sqrt{x_1^2 + x_2^2} \geq 1.2 \end{cases}$$

- Jika sebuah titik berada di dalam lingkaran, maka titik tersebut dimasukkan ke dalam kelompok 1, dan jika titik amatan berada di luar lingkaran maka dimasukkan ke dalam kelompok 2.

Langkah-langkah studi simulasinya adalah sebagai berikut :

1. Membandingkan plot hasil pengelompokan dengan metode K-Means dan Kernel K-Means.
2. Membandingkan nilai *clustering error* hasil pengelompokan dengan metode K-Means dan Kernel K-Means
3. Membandingkan nilai indeks validitas kelompok hasil pengelompokan dengan metode K-Means dan Kernel K-Means

3.2.2 Studi Kasus pada Data Riil (Data Persentase Penduduk dengan Faktor Risiko Penyebab Penyakit Hipertensi Menurut Kabupaten/Kota)

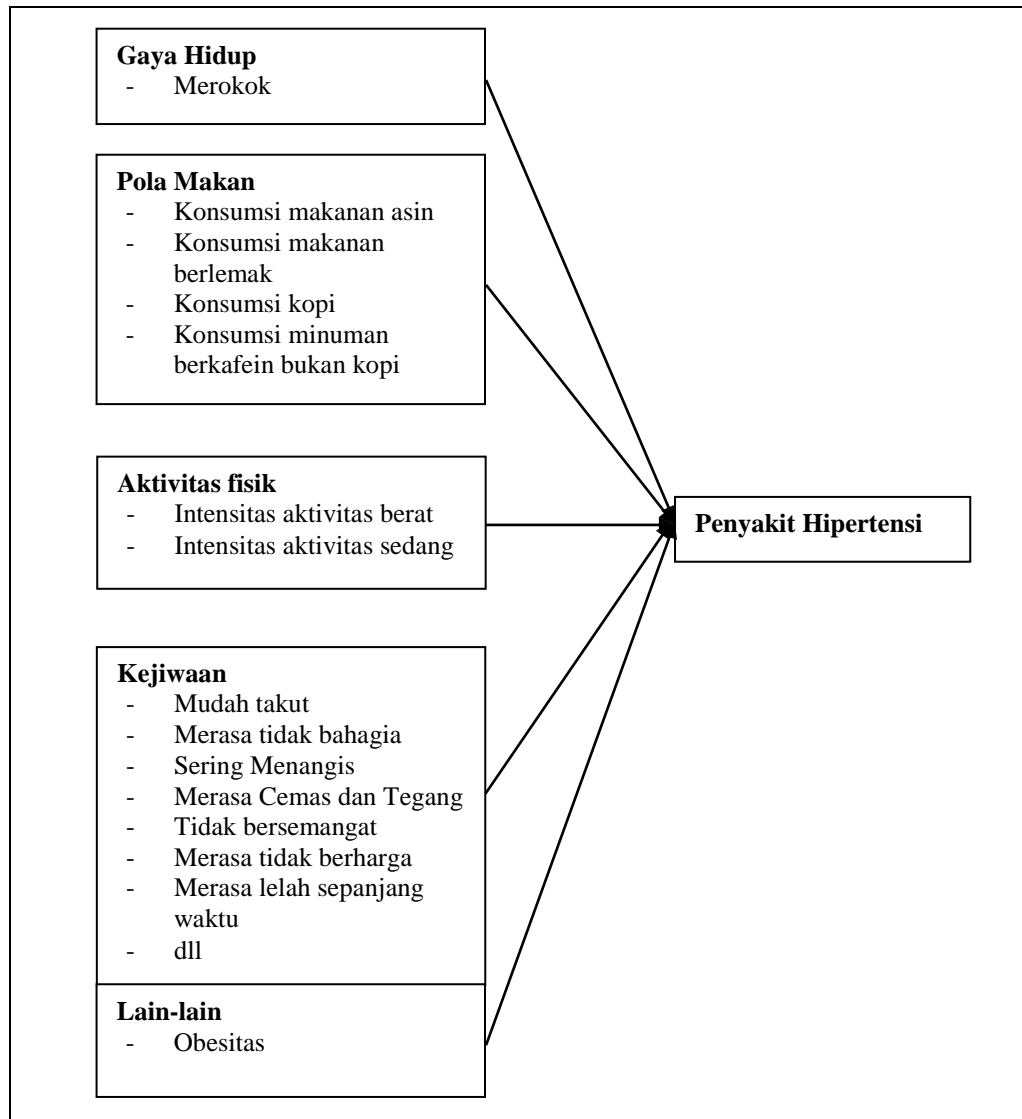
a. Sumber Data

Data riil penelitian yang digunakan adalah data Riset Kesehatan Dasar (RISKESDAS) tahun 2013 yang bersumber dari Badan Penelitian dan Pengembangan Kesehatan (Balitbangkes) Kementerian Kesehatan RI. Populasi dalam analisis ini adalah seluruh Kabupaten/Kota di Indonesia, yaitu sebanyak 497 kabupaten/kota.

b. Kerangka Konsep

Kerangka konsep yang digunakan mengacu pada penelitian Sugiharto (2007) tentang faktor-faktor risiko penyakit hipertensi. Ada banyak faktor risiko penyebab penyakit hipertensi yang termasuk dalam faktor risiko yang dapat dikendalikan (*controllable risk factor*), antara lain gaya hidup, pola makan, aktivitas fisik, kejiwaan. Faktor gaya hidup dipengaruhi oleh kebiasaan merokok. Faktor pola makan dipengaruhi oleh pola konsumsi makanan asin, berlemak/gorengan, serta konsumsi minuman yang mengandung kafein. Faktor aktifitas fisik dipengaruhi oleh intensitas aktivitas fisik. Sedangkan faktor kejiwaan dipengaruhi kondisi kejiwaan seseorang, dalam hal ini meliputi 20 variabel. Faktor lain-lain dipengaruhi oleh faktor kegemukan atau obesitas.

Faktor-faktor risiko penyebab penyakit hipertensi di atas dituangkan dalam kerangka konsep sebagaimana Gambar 3.1.



Gambar 3.1. Kerangka Konsep Faktor-Faktor Risiko Penyebab Penyakit Hipertensi

Sedangkan konsep pengelompokan wilayah berdasarkan kerangka konseptual individu mengacu pada penelitian Hatma dkk (2005) tentang kaitan pola makan dengan penyakit kronis, khususnya asupan asam lemak pada 4 (empat) suku yang berbeda di Indonesia, yaitu Suku Minangkabau, Sunda, Jawa, dan Bugis, serta penelitian Mudjianto dkk (1993) tentang perbandingan pola

konsumsi karbohidrat, protein, dan lemak pada masyarakat berpenghasilan rendah di wilayah industri dan wilayah pertanian.

Penerapan metode K-Means dan Kernel K-Means pada data riil dengan tahapan sebagai berikut :

- a. Menentukan jumlah kelompok yang optimal
- b. Mengelompokkan kabupaten/kota dengan jumlah kelompok 2 (dua) dan 3(tiga)

Adapun detail pengelompokan menggunakan metode K-means, dengan langkah-langkah sebagai berikut :

- a. Menentukan sebanyak k pusat *cluster* awal (m_k) dengan menggunakan k objek yang ditentukan secara acak
- b. Menghitung jarak masing-masing objek dengan tiap pusat cluster menggunakan jarak *Euclidean* kuadrat seperti pada persamaan (2.4)
- c. Masukkan masing-masing objek ke kelompok dengan nilai pusat cluster yang paling dekat
- d. Hitung kembali pusat cluster dari tiap kelompok yang terbentuk seperti pada persamaan (2.5)
- e. Hentikan proses jika sudah tidak ada obyek yang berpindah, jika masih ada obyek yang berpindah maka kembali mulai dari langkah b) dan seterusnya.

Untuk lebih jelasnya proses penghitungan *K-Means clustering* dapat dilihat pada Gambar 3.2

Sedangkan pengelompokan menggunakan metode Kernel K-Means langkah-langkah detailnya adalah sebagai berikut :

- a. Dengan input Kernel matriks K , banyaknya kelompok k , dan inisial kelompok (*cluster*) adalah C_1, C_2, \dots, C_k .
- b. Untuk semua nilai $n = 1, \dots, N$ dan untuk $i = 1, \dots, k$, hitung $\|\phi(X_n) - m_i\|^2$ dengan menggunakan persamaan (2.11)
- c. Tentukan nilai $C^*(X_n) = \arg \min_i \left(\|\phi(X_n) - m_i\|^2 \right)$

- d. Untuk semua kelompok, perbarui $C_i = \{X_n | C^*(X_n) = i\}$ sampai semua nilai C_i konvergen atau tidak terjadi perpindahan anggota kelompok
- e. Jika belum konvergen ulangi mulai langkah b dan seterusnya. Kriteria konvergen yang digunakan adalah
- $$C_k = \left\{ x_i \in X : \|\phi(x_i) - \phi(m_k)\|^2 \leq \|\phi(x_i) - \phi(m_h)\|^2, \forall h \neq k, h = 1, \dots, K \right\}$$

Proses penghitungan dengan metode Kernel *K-Means clustering* dapat dilihat pada Gambar 3.3

c. Variabel dan Definisi Operasional

Variabel-variabel yang digunakan dalam penelitian berdasarkan pada data hasil RISKESDAS tahun 2013 yang bersumber pada Balitbangkes Kementerian Kesehatan RI, sebagaimana disajikan dalam Tabel 3.1

Konsep dan definisi yang digunakan mengacu pada konsep dan definisi yang digunakan oleh Kemenkes RI dan BPS, yaitu:

1. Sering menderita sakit kepala.
Persentase penduduk usia 15 tahun ke atas yang mengalami gangguan sakit kepala.
2. Tidak nafsu makan
Persentase penduduk usia 15 tahun ke atas yang mengalami gangguan tidak nafsu makan.
3. Sulit tidur.
Persentase penduduk usia 15 tahun ke atas yang mengalami gangguan sulit tidur.
4. Suasana mudah takut
Persentase penduduk usia 15 tahun ke atas yang sering mudah takut.
5. Merasa tegang, cemas, kuatir
Persentase penduduk usia 15 tahun ke atas yang sering merasa tegang, cemas, kuatir.
6. Merasa tangan gemetar
Persentase penduduk usia 15 tahun ke atas yang sering merasa tangan gemetar.

7. Pencernakan terganggu/buruk
Persentase penduduk usia 15 tahun ke atas yang mengalami gangguan pencernaan.
8. Sulit untuk berpikir jernih
Persentase penduduk usia 15 tahun ke atas yang sering sulit berfikir jernih.
9. Merasa tidak bahagia
Persentase penduduk usia 15 tahun ke atas yang merasa tidak bahagia.
10. Menangis lebih sering
Persentase penduduk usia 15 tahun ke atas yang sering menangis.
11. Merasa sulit menikmati kegiatan sehari-hari
Persentase penduduk usia 15 tahun ke atas yang merasa sulit menikmati kegiatan sehari-hari.
12. Sulit untuk mengambil keputusan
Persentase penduduk usia 15 tahun ke atas yang merasa sulit untuk mengambil keputusan.
13. Pekerjaan sehari-hari terganggu
Persentase penduduk usia 15 tahun ke atas yang mengalami pekerjaan sehari-hari.
14. Tidak mampu melaksanakan hal-hal yang bermanfaat dalam hidup
Persentase penduduk usia 15 tahun ke atas yang merasa tidak mampu melaksanakan hal-hal yang bermanfaat dalam hidup.
15. Kehilangan minat pada berbagai hal
Persentase penduduk usia 15 tahun ke atas yang kehilangan minat pada berbagai hal.
16. Merasa tidak bahagia
Persentase penduduk usia 15 tahun ke atas yang merasa tidak bahagia.
17. Mempunyai pikiran untuk mengakhiri hidup
Persentase penduduk usia 15 tahun ke atas yang mempunyai pikiran untuk mengakhiri hidup.
18. Merasa lelah sepanjang waktu
Persentase penduduk usia 15 tahun ke atas yang merasa lelah sepanjang waktu.

19. Mengalami rasa tidak enak di perut

Persentase penduduk usia 15 tahun ke atas yang mengalami rasa tidak enak di perut.

20. Mudah lelah

Persentase penduduk usia 15 tahun ke atas yang mudah lelah.

21. Merokok

Persentase penduduk usia 15 tahun ke atas yang merokok. Merokok adalah cara utama tembakau digunakan di seluruh dunia, dan rokok filter produksi pabrik menjadi semakin dominan sebagai produk tembakau utama.

22. Melakukan aktivitas berat

Persentase penduduk usia 15 tahun ke atas yang melakukan aktivitas berat terus-menerus minimal 10 menit setiap kali melakukannya. Aktivitas fisik berat merupakan aktivitas fisik yang memerlukan kerja fisik berat dan menyebabkan nafas atau denyut nadi meningkat cepat.

23. Melakukan aktivitas sedang

Persentase penduduk usia 15 tahun ke atas yang melakukan aktivitas sedang terus-menerus minimal 10 menit. Aktivitas fisik sedang merupakan aktivitas fisik yang memerlukan kerja fisik sedang dan sedikit peningkatan denyut nadi atau nafas

24. Konsumsi makanan asin

Persentase Penduduk 15 tahun ke atas yang mengkonsumsi makanan/minuman asin lebih dari 1 kali dalam sehari. Hal ini mengacu pada batas konsumsi gula, garam, dan lemak yang disarankan oleh Kementerian Kesehatan RI per orang per hari, yaitu 50 gram (4 sendok makan) gula, 2000 miligram natrium/sodium atau 5 gram garam (1 sendok teh), dan untuk lemak hanya 67 gram (5 sendok makan minyak).

25. Konsumsi makanan berlemak/berkolesterol/gorengan

Persentase Penduduk 15 tahun ke atas yang mengkonsumsi makanan/minuman berlemak lebih dari 1 kali dalam sehari.

26. Konsumsi kopi

Persentase Penduduk 15 tahun ke atas yang mengkonsumsi kopi lebih dari 1 kali dalam sehari.

27. Konsumsi minuman berkafein buatan bukan kopi

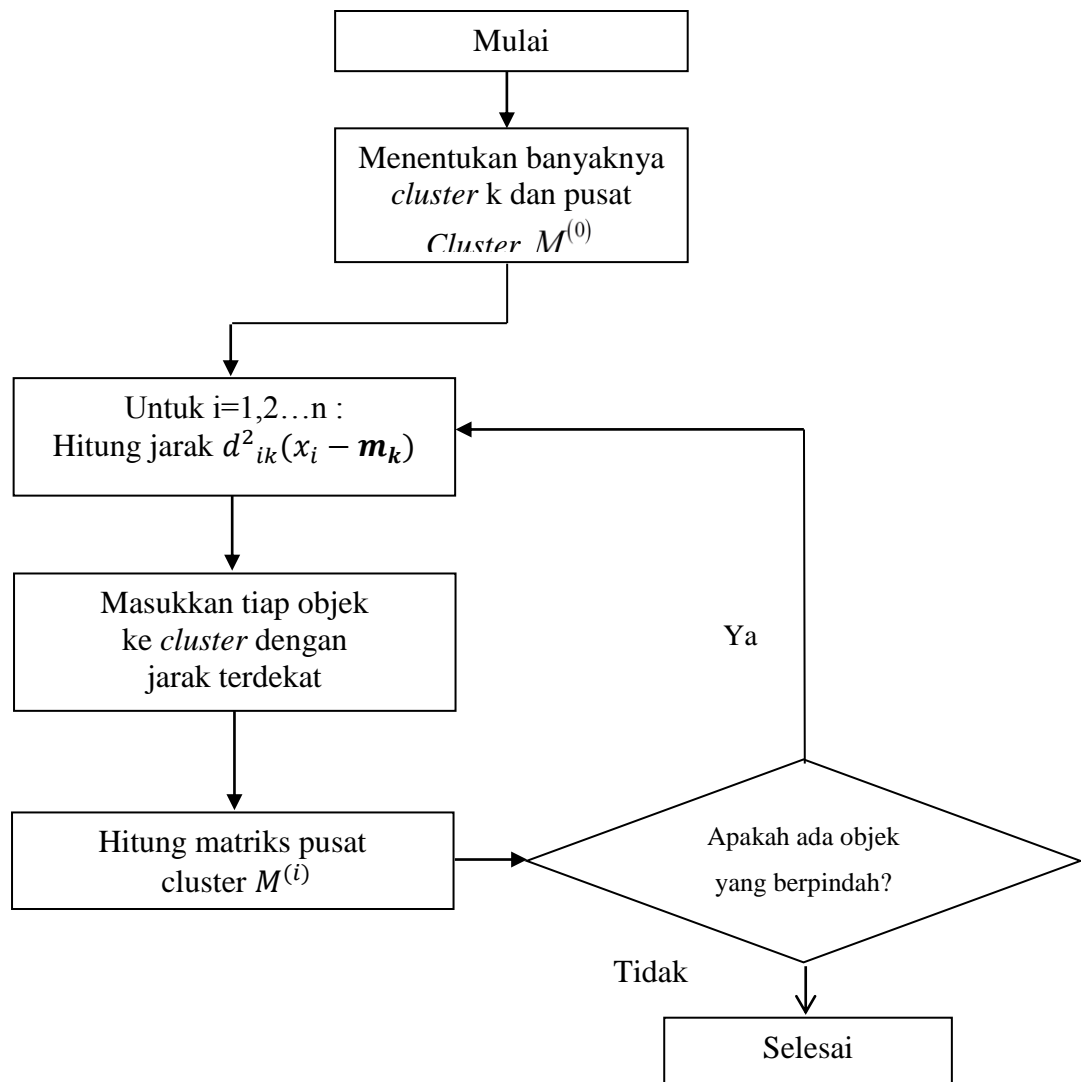
Persentase Penduduk 15 tahun ke atas yang mengkonsumsi makanan/minuman berkafein lebih dari 1 kali dalam sehari.

28. Obesitas

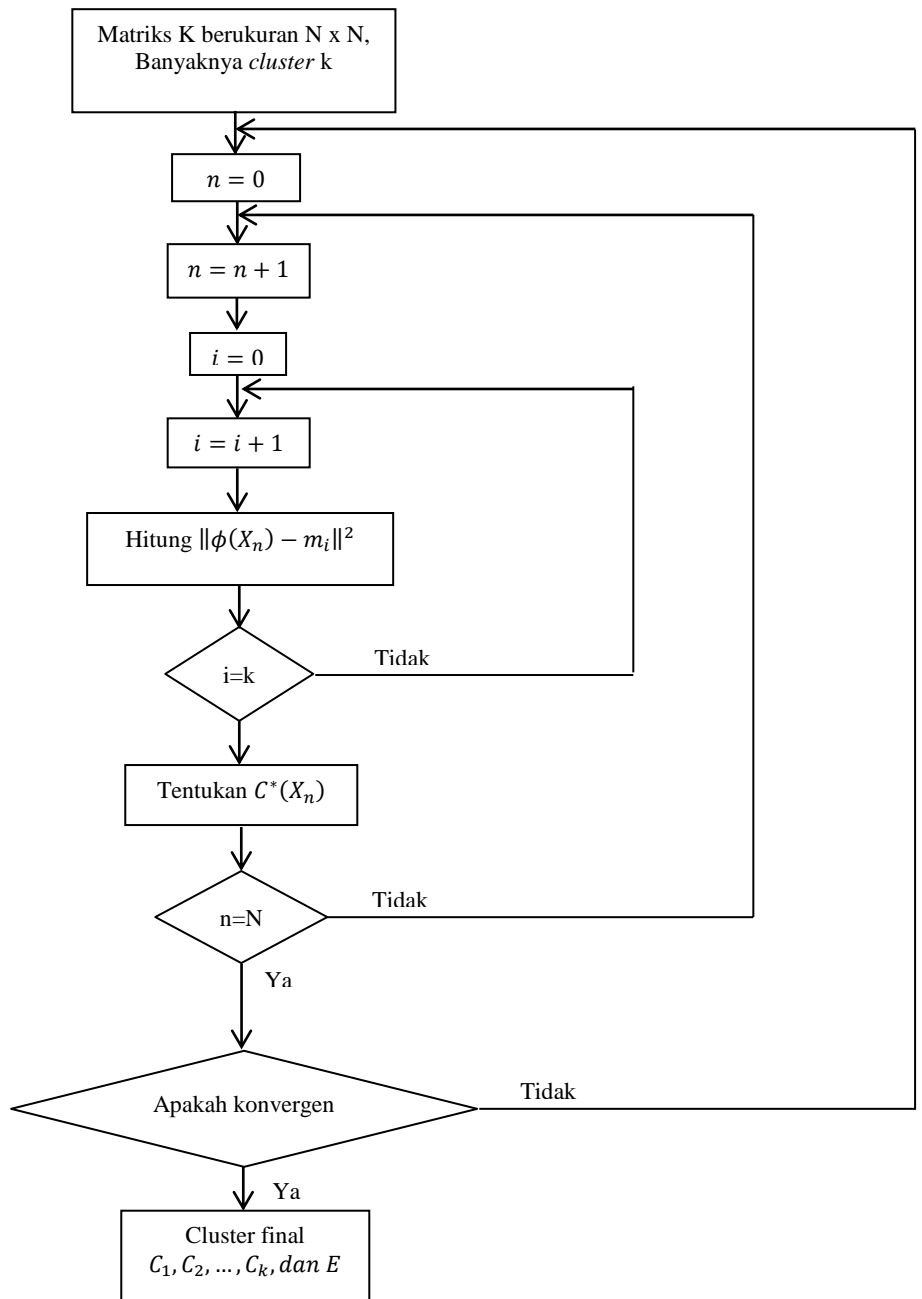
Persentase penduduk 15 tahun ke atas yang memiliki Indeks Massa Tubuh (IMT) lebih dari 25. IMT adalah berat badan (kg) dibagi tinggi badan kuadrat (m^2).

Tabel. 3.1 Variabel dan Label yang Digunakan dalam Penelitian

Variabel	Label
X_1	Sering menderita sakit kepala
X_2	Tidak nafsu makan
X_3	Sulit tidur
X_4	Suasana mudah takut
X_5	Merasa tegang, cemas, khawatir
X_6	Merasa tangan gemetar
X_7	Pencernakan terganggu/buruk
X_8	Sulit untuk berfikir jernih
X_9	Merasa tidak bahagia
X_{10}	Mengangis lebih sering
X_{11}	Merasa sulit menikmati kegiatan sehari-hari
X_{12}	Sulit untuk mengambil keputusan
X_{13}	Pekerjaan sehari-hari terganggu
X_{14}	Tidak mampu melakukan hal-hal yang bermanfaat dalam hidup
X_{15}	Kehilangan minat pada berbagai hal
X_{16}	Merasa tidak berharga
X_{17}	Mempunyai pikiran untuk mengakhiri hidup
X_{18}	Merasa lelah sepanjang waktu
X_{19}	Mengalami tidak enak di perut
X_{20}	Mudah lelah
X_{21}	Merokok
X_{22}	Melakukan aktivitas berat
X_{23}	Melakukan aktivitas sedang
X_{24}	Konsumsi makanan asin
X_{25}	Konsumsi makanan berlemak/berkolesterol/gorengan
X_{26}	Konsumsi kopi
X_{27}	Konsumsi minuman berkafein buatan bukan kopi
X_{28}	Obesitas



Gambar 3.2 Alur Kerja Metode *K-Means Clustering*



Gambar 3.3 Alur Kerja Metode *Kernel K-Means Clustering*

BAB IV HASIL DAN PEMBAHASAN

Pada Bab 4 ini akan dibahas mengenai analisis dan pembahasan dari tujuan penelitian, yaitu mengkaji pembentukan cluster (kelompok) dengan menggunakan metode Kernel K-Means (KKM), dan implementasi metode KKM pada data penelitian untuk kasus pengelompokan kabupaten/kota berdasarkan penduduk dengan faktor risiko penyebab penyakit hipertensi.

4.1 *Kernel K Means Clustering*

4.1.1 Sifat-Sifat Kernel

Fungsi kernel secara umum memiliki sifat sebagai berikut (dikenal sebagai *Theorema Mercer*) :

- a. Matriks kernel adalah berupa matrik simetris, dimana
$$K(\mathbf{x}_i, \mathbf{x}_j) = K(\mathbf{x}_j, \mathbf{x}_i)$$
- b. Matriks $N \times N$ dengan elemen matriks $K(\mathbf{x}_i, \mathbf{x}_j) \forall i, j = 1, \dots, N$ merupakan matriks *positif semi definite*.

Sedangkan sifat lain dari Kernel untuk sembarang ruang sampel X dengan kernel k_1 dan k_2 yang melalui X , antara lain adalah :

- c. $k(x, y) = k_1(x, y) + k_2(x, y)$
- d. $k(x, y) = ak_1(x, y)$, dimana $a > 0$
- e. $k(x, y) = f(x) \cdot f(y)$ untuk sembarang fungsi f terhadap x
- f. $k(x, y) = k_1(x, y) \cdot k_2(x, y)$
- g. $k(x, y) = \frac{k_1(x, y)}{\sqrt{k_1(x, x)} \sqrt{k_2(y, y)}}$

Pembuktian dari sifat-sifat kernel di atas adalah sebagai berikut :

- a. Matriks kernel \mathbf{K} atau disebut juga *Gram Matrix* \mathbf{K} adalah berupa matriks simetris yang elemennya merupakan hasil dari pemetaan datapoint pada set data $\{x_i\}_{i=1}^N$

$$\mathbf{K} = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_N) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_N, x_1) & k(x_N, x_2) & \cdots & k(x_N, x_N) \end{bmatrix}$$

Dimana matriks kernel adalah matriks bujur sangkar dan simetris, sehingga $k(x_i, x_j) = k(x_j, x_i)$.

- b. Matriks $N \times N$ dengan elemen matriks $K(\mathbf{x}_i, \mathbf{x}_j) \forall i, j = 1, \dots, N$ merupakan matriks *positif semi definite*.

Matriks \mathbf{A} dikatakan *positive semi definite* (PSD) jika nilai *eigen valuenya non-negative* ($\lambda_i(\mathbf{A}) \geq 0$) untuk semua $x \in \mathcal{X}$ harus memenuhi $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$.

Sedangkan pada matriks kernel, untuk semua nilai $i, j \in 1, \dots, N$

$$K_{i,j} = k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle_F$$

Dengan demikian untuk $\mathbf{v} \in R^N$:

$$\begin{aligned} \mathbf{v}^T \mathbf{K} \mathbf{v} &= \sum_{i,j=1}^N v_i K_{i,j} v_j = \sum_{i,j=1}^N v_i \langle \phi(x_i), \phi(x_j) \rangle_F v_j \\ &= \left\langle \sum_{i=1}^N v_i \phi(x_i), \sum_{j=1}^N v_j \phi(x_j) \right\rangle_F \\ &= \left\| \sum_{i=1}^N v_i \phi(x_i) \right\|_F^2 \geq 0 \end{aligned}$$

- c. $k(x, y) = k_1(x, y) + k_2(x, y)$

Representasi $k(x, y)$ dalam ruang *feature (feature space)* adalah

$$\langle \phi(x), \phi(y) \rangle, \text{ dimana } \phi(x) = \begin{bmatrix} \phi_1(x) \\ \phi_2(x) \end{bmatrix} \text{ dan } \phi(y) = \begin{bmatrix} \phi_1(y) \\ \phi_2(y) \end{bmatrix}, \text{ sehingga}$$

$$\begin{aligned}
k(x, y) &= \langle \phi(x), \phi(y) \rangle \\
&= \left\langle \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix}, \begin{pmatrix} \phi_1(y) \\ \phi_2(y) \end{pmatrix} \right\rangle \\
&= \langle \phi_1(x), \phi_1(y) \rangle + \langle \phi_2(x), \phi_2(y) \rangle \\
&= k_1(x, y) + k_2(x, y)
\end{aligned}$$

d. $k(x, y) = ak_1(x, y)$, dimana $a > 0$

Telah disebutkan di atas bahwa $k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle_F$, sehingga :

$$\begin{aligned}
k(x, y) &= (\sqrt{a}\phi_1^1(x), \dots, \sqrt{a}\phi_N^1(x))(\sqrt{a}\phi_1^1(y), \dots, \sqrt{a}\phi_N^1(y)) \\
&= ak_1(x, y)
\end{aligned}$$

e. $k(x, y) = f(x) \cdot f(y)$ untuk sembarang fungsi f terhadap x

Hanya ada satu *feature* yang didefinisikan oleh $f()$.

f. $k(x, y) = k_1(x, y) \cdot k_2(x, y)$

$$\begin{aligned}
k(x, y) &= \langle \phi(x), \phi(y) \rangle \\
&= \sum_{i,j=1}^N \phi(x)_{i,j} \phi(y)_{i,j} \\
&= \sum_{i,j=1}^N \phi_1(x)_i \phi_2(x)_j \phi_1(y)_i \phi_2(y)_j \\
&= \left(\sum_{i=1}^N \phi_1(x)_i \phi_1(y)_i \right) \left(\sum_{j=1}^N \phi_2(x)_j \phi_2(y)_j \right) \\
&= \langle \phi_1(x), \phi_1(y) \rangle \langle \phi_2(x), \phi_2(y) \rangle \\
&= k_1(x, y) \cdot k_2(x, y)
\end{aligned}$$

g. $k(x, y) = \frac{k_1(x, y)}{\sqrt{k_1(x, x)} \sqrt{k_2(y, y)}}$

Norm dari $\phi(x)$ adalah $\|\phi(x)\|$, dimana :

$$\|\phi(x)\| = \sqrt{\langle \phi(x), \phi(x) \rangle} = \sqrt{K(x, x)}$$

Didefinisikan bahwa $\phi(x) = \frac{\phi(x)}{\|\phi(x)\|}$, sehingga :

$$\begin{aligned}
k(x, y) &= \langle \phi(x), \phi(y) \rangle \\
&= \frac{k(x, y)}{\|\phi(x)\| \|\phi(y)\|} = \frac{k(x, y)}{\sqrt{k(x, x)} \sqrt{k(y, y)}}
\end{aligned}$$

4.1.2 Algoritma Kernel K-Means *Clustering*

Kernel K-Means merupakan pengembangan dari algoritma K-Means, dengan cara memetakan data ke *space* baru yang berdimensi tinggi menggunakan metode kernel untuk meningkatkan akurasi hasil pengelompokan. Dengan menggunakan metode Kernel K-Means diharapkan data dapat dipisahkan dengan lebih baik, karena data yang *overlap* atau non linier dapat menjadi linier di ruang berdimensi tinggi. Fungsi kernel yang biasa digunakan adalah kernel Gaussian. Algoritma Kernel K-Means *Clustering* sebagaimana disajikan pada Gambar 3.3 adalah sebagai berikut :

Dengan input adalah matriks kernel K , banyaknya kelompok k , dan inisial cluster C_k

1. Inisialisasi

Menentukan banyaknya kelompok, $2 \leq K < n$; kemudian menentukan pusat kelompok secara random (m_k). Pada tahap awal setiap observasi dimasukkan ke pusat kelompok terdekat.

2. Tahap Representasi

Menghitung kembali pusat kelompok m_k dengan menggunakan persamaan (4.3)

3. Tahap Alokasi

Untuk $i = 1$ sampai dengan n , hitung $\|\phi(X_n) - \phi(m_k)\|^2$. Pada tahap ini pusat kelompok m_k dianggap *fixed*. Kelompok $C_k (k = 1, \dots, K)$ yang meminimalkan kriteria pengelompokan ε_K^ϕ diperbaharui (*diupdate*) sehingga memenuhi aturan berikut :

$$C_k = \left\{ x_i \in X : \|\phi(x_i) - \phi(m_k)\|^2 \leq \|\phi(x_i) - \phi(m_h)\|^2, \forall h \neq k, h = 1, \dots, K \right\}$$

4. Hentikan proses jika sudah konvergen, dimana tidak terjadi perpindahan anggota kelompok. Jika proses belum konvergen, ulangi mulai langkah (2).

Prinsip dalam KKM adalah meminimalkan nilai \mathcal{E}_K^ϕ , dimana

$$\mathcal{E}_K^\phi = \sum_{n=1}^N \sum_{k=1}^K z_{kn} \left\| \phi(X_n) - \phi(\mathbf{m}_k) \right\|^2. \text{ Dengan mensubstitusi persamaan (2.12) dan}$$

(2.13) ke dalam persamaan (2.11) diperoleh persamaan berikut ini :

$$\begin{aligned} \mathcal{E}_K^\phi &= \sum_{n=1}^N \sum_{k=1}^K z_{kn} \left\| \phi(X_n) - \phi(\mathbf{m}_k) \right\|^2 \\ &= \sum_{n=1}^N \sum_{k=1}^K z_{kn} \left(\phi(X_n) - \phi(\mathbf{m}_k) \right)^T \left(\phi(X_n) - \phi(\mathbf{m}_k) \right) \\ &= \sum_{n=1}^N \sum_{k=1}^K z_{kn} \left(\begin{aligned} &\phi(X_n)^T \phi(X_n) - \\ &\frac{2}{N_k} \sum_{m=1}^N z_{km} \phi(X_m)^T \phi(X_n) + \\ &\frac{1}{N_k^2} \sum_{p=1}^N \sum_{l=1}^N z_{kp} z_{kl} \phi(X_p)^T \phi(X_l) \end{aligned} \right) \\ &= \sum_{n=1}^N \sum_{k=1}^K z_{kn} \left(\begin{aligned} &K(X_n, X_n) - \\ &\frac{2}{N_k} \sum_{m=1}^N z_{km} K(X_m, X_n) + \\ &\frac{1}{N_k^2} \sum_{p=1}^N \sum_{l=1}^N z_{kp} z_{kl} K(X_p, X_l) \end{aligned} \right) \end{aligned}$$

Jika jarak antara data ke n dengan pusat *cluster* ke-k dalam ruang berdimensi tinggi dilambangkan dengan δ_{kn} , maka:

$$\begin{aligned} \delta_{kn} &= \left\| \phi(X_n) - \phi(\mathbf{m}_k) \right\|^2 \\ &= \left(\begin{aligned} &K(X_n, X_n) - \\ &\frac{2}{N_k} \sum_{m=1}^N z_{km} K(X_m, X_n) + \\ &\frac{1}{N_k^2} \sum_{p=1}^N \sum_{l=1}^N z_{kp} z_{kl} K(X_p, X_l) \end{aligned} \right) \end{aligned}$$

Sehingga $\varepsilon_K^\phi = \sum_{n=1}^N \sum_{k=1}^K z_{kn} \delta_{kn}$

Dengan menggunakan *kernel trick*, memungkinkan untuk menghitung jarak *euclidean* pada ruang *feature* sebagai berikut:

$$\begin{aligned} \|\phi(X_n) - \phi(\mathbf{m}_k)\|^2 &= (\phi(X_n) - \phi(\mathbf{m}_k))^T (\phi(X_n) - \phi(\mathbf{m}_k)) \\ &= \phi(X_n)^T \phi(X_n) - 2\phi(X_n)^T \phi(\mathbf{m}_k) + \phi(\mathbf{m}_k)^T \phi(\mathbf{m}_k) \quad (4.1) \\ &= K(X_n, X_n) - 2K(X_n, \mathbf{m}_k) + K(\mathbf{m}_k, \mathbf{m}_k) \end{aligned}$$

Pada kernel Gaussian, $K^g(X_i, X_i) = 1 \quad \forall i$, sehingga persamaan (4.1) menjadi :

$$\|\phi(X_n) - \phi(\mathbf{m}_k)\|^2 = 2(1 - K(X_n, \mathbf{m}_k)) \quad (4.2)$$

Dan $\varepsilon_K^\phi = \sum_{n=1}^N \sum_{k=1}^K z_{kn} (2(1 - K(X_n, \mathbf{m}_k)))$

Kemudian untuk memperoleh nilai pusat kelompok (\mathbf{m}_k) pada ruang *feature* dengan cara sebagai berikut :

$$\begin{aligned} \nabla \mathbf{m}_k \varepsilon_K^\phi &= 0 \\ \sum_{n=1}^N z_{kn} \nabla \mathbf{m}_k (\|\phi(X_n) - \phi(\mathbf{m}_k)\|^2) &= 0 \\ \sum_{n=1}^N z_{kn} \nabla \mathbf{m}_k (2 - 2K(X_n, \mathbf{m}_k)) &= 0 \\ \sum_{n=1}^N z_{kn} \nabla \mathbf{m}_k (e^{-\|X_n - \mathbf{m}_k\|^2 / \sigma^2}) &= 0 \\ \sum_{n=1}^N z_{kn} \left(e^{-\|X_n - \mathbf{m}_k\|^2 / \sigma^2} \right) \left(\frac{2}{\sigma^2} (X_n - \mathbf{m}_k) \right) &= 0 \\ v_i \sum_{n=1}^N z_{kn} K(X_n, \mathbf{m}_k) &= \sum_{n=1}^N z_{kn} K(X_n, \mathbf{m}_k) X_n \\ \mathbf{m}_k &= \frac{\sum_{n=1}^N z_{kn} K(X_n, \mathbf{m}_k) X_n}{\sum_{n=1}^N z_{kn} K(X_n, \mathbf{m}_k)} \quad (4.3) \end{aligned}$$

4.1.3 Kompleksitas Algoritma Kernel K Means

Untuk menganalisa efisiensi sebuah algoritma, dalam hal ini terkait kompleksitas, yang perlu dilakukan adalah perhitungan pertumbuhan fungsi. Fungsi pertumbuhan biasa ditulis dengan notasi asimtotik. Terdapat beberapa

jenis notasi asimtotik, diantaranya adalah notasi Big-O, biasa disebut juga Notasi Landau.

Definisi Big- O adalah , misalkan f dan g adalah dua fungsi yang terdefiniskan pada bilangan real x . Fungsi $f(x)$ dikatakan *Most of order* dari $g(x)$, $x \rightarrow \infty$, ditulis dengan simbol :

$$f(x) = O(g(x)), x \rightarrow \infty$$

Jika terdapat suatu konstanta $M < \infty$, sehingga $\frac{|f(x)|}{|g(x)|} < M$

Atau jika x adalah suatu ukuran atau besarnya pengamatan, suatu algoritma mempunyai order $O(g(x))$ jika, untuk $x \rightarrow \infty$ banyaknya perhitungan $\rightarrow Mg(x)$, dimana M adalah konstanta yang tidak tergantung pada nilai x .

Dalam Ferreira dkk (2014), disebutkan bahwa kompleksitas dari algoritma KKM untuk single iterasi adalah $O(nKp)$, yang berarti untuk menyelesaikan perhitungan dengan metode KKM diperlukan sebanyak nKp langkah, dengan n adalah banyaknya observasi, K adalah banyaknya kelompok, dan p adalah banyaknya variabel.

4.2 Perbandingan Pengelompokan metode K-Means dan Kernel K-Means pada Data Simulasi

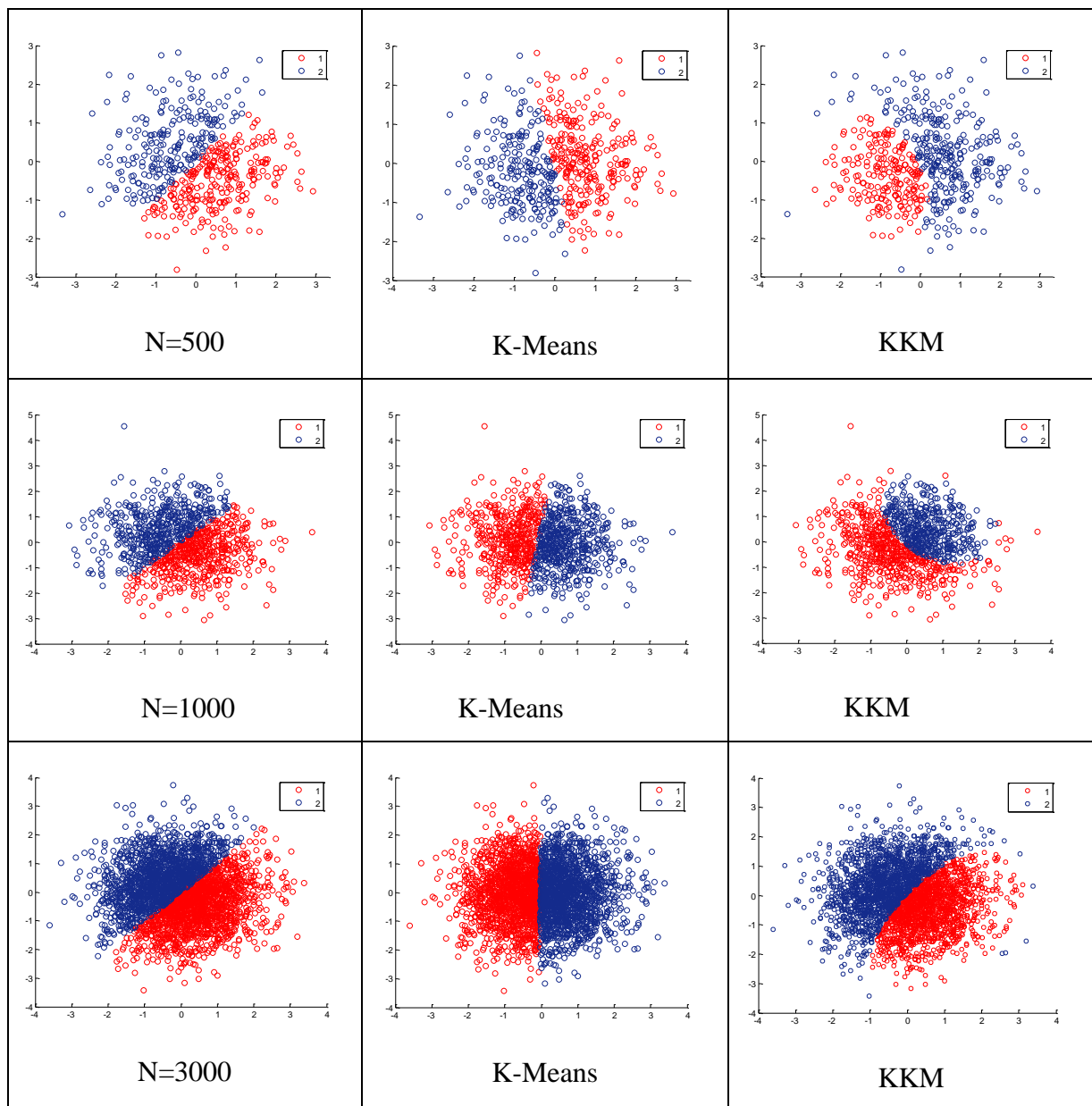
Tahap awal dari algoritma K-Means menurut Girolami (2002) adalah membuat inisiasi, yaitu menentukan banyaknya kelompok (K), dan menentukan pusat kelompok (m) secara random. Kemudian menghitung jarak antara objek terhadap pusat kelompok dengan menggunakan formula jarak Euclidean, dan setiap objek kembali dikelompokkan menurut jarak terdekat terhadap pusat kelompok yang baru terbentuk. Hitung kembali pusat kelompok dari kelompok yang terbentuk, ulangi lagi menghitung jarak setiap objek terhadap pusat kelompok sampai tidak terjadi perubahan anggota kelompok dalam tiap kelompok (cluster).

Sedangkan pada pengelompokan dengan metode Kernel K-Means, sebagai inputnya adalah matriks kernel hasil pemetaan data ke dalam ruang berdimensi tinggi, untuk selanjutnya dilakukan langkah sebagaimana pengelompokan dengan metode K-Means, yaitu menentukan pusat kelompok dan jumlah kelompok, kemudian menghitung jarak antara elemen matrik kernel dengan pusat cluster, sampai konvergen sebagaimana pada metode K-Means.

Untuk mengetahui perbedaan pengelompokan dengan menggunakan metode K-Means dan Kernel K Means, terlebih dahulu dilakukan perbandingan plot data simulasi yang terbagi dalam 2 kelompok dengan menggunakan kedua metode pengelompokan di atas. Data simulasi yang digunakan bersumber pada data penelitian sebelumnya (data penelitian Epa Suryanto, 2015), yang berupa data linier dan non linier, dengan jumlah observasi masing-masing 500, 1000, dan 3000.

Data simulasi berupa data yang dipisahkan secara linier dan lingkaran (non linier) dibangkitkan dari distribusi normal dengan rata-rata 0 dan varians 1. Variabel prediktor terdiri dari dua variabel, x_1 dan x_2 . Selanjutnya variabel prediktor dikategorikan menjadi dua kelas, yaitu kelas $\{1\}$ dan $\{2\}$. Pada data linier, apabila nilai x_1 lebih besar dari x_2 maka data tersebut berada di kelas $\{1\}$, sedangkan jika x_1 kurang dari x_2 maka data tersebut berada di kelas $\{2\}$. Pada data non linier, bidang pemisah berupa lingkaran dengan jari-jari tertentu. Apabila sebuah titik berada di dalam lingkaran, maka titik tersebut dikategorikan sebagai kelas $\{1\}$, dan apabila titik tersebut berada di luar lingkaran, maka dimasukkan ke dalam kelas $\{2\}$.

Plot hasil pengelompokan dengan metode KM dan KKM yang lebih menyerupai plot data simulasi menunjukkan metode yang lebih baik. Selain membandingkan plot hasil pengelompokan dengan kedua metode tersebut, dapat juga dengan melihat nilai e (jarak minimum antara pusat kelompok ke- k dengan titik observasi ke- n .)



Sumber : Hasil olah data

Gambar 4.1 Perbedaan Plot data Linier Hasil Pengelompokan Metode KM dan KKM

Plot data linier Gambar 4.1 menunjukkan kedua kelompok dipisahkan oleh pemisah berupa garis lurus. Titik-titik dengan warna merah menunjukkan titik berada pada kelompok 1, sedangkan titik dengan warna biru menunjukkan titik berada pada kelompok 2. Semakin banyak jumlah data semakin terlihat jelas batas pemisah antara kedua kelompok. Sesuai dengan fungsi untuk membangkitkan data linier pada pembahasan sebelumnya, bahwa jika nilai $x_1 > x_2$, maka data

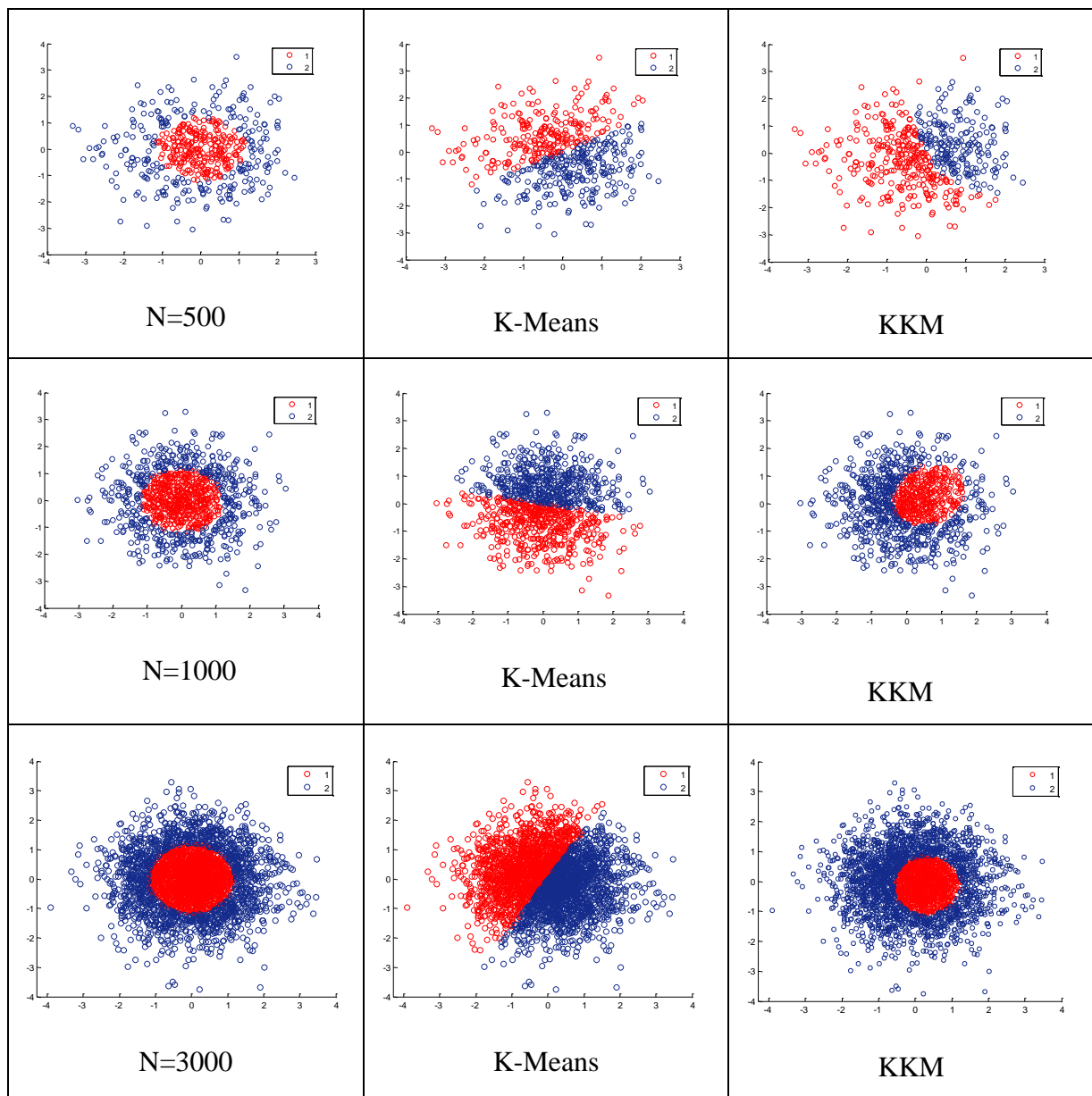
dimasukkan ke dalam kelompok 1, dan jika $x_1 \leq x_2$ maka data dimasukkan dalam kelompok 2.

Berdasarkan Gambar 4.1, terlihat bahwa untuk pengelompokan pada data yang terpisah secara linier untuk jumlah data 500 dan 1000 bentuk plot yang lebih menyerupai adalah pengelompokan dengan metode K-Means. Sedangkan pada jumlah pengamatan 3000, baik metode K-means maupun Kernel K-Means plotnya menyerupai data asli. Dengan semakin banyak data yang digunakan plot yang terbentuk lebih menyerupai data asli.

Sedangkan pengelompokan pada data yang bersifat *non-linearly separable*, plot yang lebih menyerupai data asli adalah plot pengelompokan dengan metode Kernel K-Means (Gambar 4.2).

Gambar 4.2 menunjukkan pemisah antar kelompok yang berbentuk lingkaran dan berpusat di titik (0,0). Plot yang ditunjukkan memperlihatkan bahwa jika suatu titik berada di dalam lingkaran, maka akan dimasukkan ke dalam kelompok (1), dan apabila di luar lingkaran akan dimasukkan ke dalam kelompok (2). Plot dengan dengan bidang pemisah berupa lingkaran dengan titik yang berwarna merah menunjukkan titik data yang berada pada kelompok 1, sedangkan titik-titik dengan warna biru menunjukkan data yang berada pada kelompok 2. Dengan semakin banyak data yang digunakan memperlihatkan semakin jelas bidang pemisah yang digunakan.

Berdasarkan Gambar 4.2, terlihat bahwa pengelompokan pada data yang memiliki pemisah non linier dengan $n=1000$ dan 3000 lebih menyerupai data asli.



Sumber : Hasil olah data

Gambar 4.2 Perbedaan Plot data Non Linier Hasil Pengelompokan Metode KM dan KKM

Selain dengan membandingkan plot data, untuk melihat perbandingan pengelompokan antara metode KM dan KKM dapat dengan melihat nilai *clustering error* (e) yang lebih kecil. Nilai e dalam hal ini adalah jarak antara titik observasi ke- n dengan jarak titik pusat cluster ke- k . Jarak yang minimum menunjukkan observasi dalam kelompok lebih homogen. Di bawah ini disajikan

tabel perbedaan nilai *clustering error* pengelompokan data simulasi dengan menggunakan metode KM dan KKM.

Tabel 4.1 Perbedaan Rata-rata Nilai *Clustering Error* (e) Pengelompokan dengan Metode KM dan KKM

Jenis Data Simulasi	Metode KM (e)	Metode KKM (e)
Linier (n=500)	1.27	0.71
Linier (n=1000)	1.35	0.70
Linier (n=3000)	1.36	0.73
Non Linier (n=500)	1.47	0.68
Non Linier (n=1000)	1.34	0.78
Non Linier (n=3000)	1.33	0.76

Sumber : Hasil Olah Data

Berdasarkan Tabel 4.1, metode KKM menghasilkan nilai *clustering error* yang lebih kecil daripada metode KM, yang berarti keseragaman dalam tiap kelompok yang dibentuk dengan menggunakan metode KKM lebih tinggi dibandingkan pengelompokan dengan metode KM. Sehingga dalam hal ini metode KKM lebih baik dibandingkan dengan metode KM. Akibat dari penentuan pusat kelompok (pusat *cluster*) secara random sebagaimana disebutkan di awal, mengakibatkan nilai besaran *clustering error* maupun besaran pusat kelompok yang berubah-ubah.

Selain dengan melihat plot dan nilai *clustering error*, untuk melihat metode pengelompokan mana yang lebih baik performansinya, dapat juga dengan melihat nilai validitas kelompok. Dengan menggunakan bantuan paket program Matlab dan R versi 2.15.3 dilakukan pengujian tingkat akurasi pengelompokan untuk mengukur keterbandingan kualitas pengelompokan dengan menggunakan metode KKM dan KM dengan pendekatan ukuran validitas kelompok dengan menggunakan data simulasi sebagaimana tersebut di atas.

Ukuran validitas kelompok digunakan untuk mengukur perbandingan kualitas metode *clustering*. Indeks validitas yang digunakan dalam penelitian ini adalah indeks validitas internal. Indeks Validitas internal yang digunakan dalam penelitian ini mengacu pada penelitian Liu (2010), dimana berdasarkan hasil penelitian Liu dkk, indeks S_{Dbw} merupakan indeks yang memiliki performansi

paling baik dibandingkan ke-sepuluh indeks lainnya sebagaimana disajikan dalam Tabel 2.1 Rekapitulasi hasil pengolahan terhadap ukuran validitas kelompok tersaji dalam tabel berikut ini.

Tabel 4.2 Indeks Validitas Kelompok pada Metode KM dan KKM pada Data Simulasi dan Data Penelitian

Metode	Kernel yang digunakan	Linier			Nonlinier			Data Penelitian
		Sim 1	Sim 2	Sim 3	Sim 1	Sim 2	Sim 3	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
KM	Gaussian	1.909	1.919	2.047	2.034	2.070	2.109	Inf
	Polynomial	1.887	1.960	2.047	2.139	2.000	2.093	Inf
	Linier	1.887	1.960	2.047	2.139	2.000	2.093	Inf
KKM	Gaussian	1.797	1.990	2.007	1.864	1.864	1.877	1.843
	Polynomial	1.752	1.805	1.795	1.737	1.767	1.908	1.339
	Linier	1.752	1.805	1.795	1.737	1.767	1.908	1.339

Sumber : Hasil Olah Data

Tabel di atas menggambarkan hasil penghitungan nilai indeks validitas kelompok terhadap data simulasi untuk data linier dan non linier dengan tiga populasi yang berbeda dengan menggunakan tiga jenis kernel, yaitu kernel gaussian, kernel polynomial, dan kernel linier. Populasi 1 menunjukkan observasi terhadap populasi dengan n sebanyak 500, sedangkan populasi 2 dan 3 masing – masing dengan n berturut-turut sebanyak 1000 dan 3000. Indeks S_{Dbw} mengukur varians di dalam kelompok (*intra cluster variance*) dan varian antar kelompok (*inter cluster variance*), sebagaimana dijelaskan dalam persamaan (2.23). Nilai indeks S_{Dbw} yang diharapkan adalah nilai yang terkecil. Berdasarkan tabel rekapitulasi di atas, pada data linier, untuk data populasi 1 dan 3, pada semua jenis kernel menunjukkan indeks S_{Dbw} pada metode KKM lebih kecil dibanding pada metode KM. Pada data linier dengan sampel 1000 indeks S_{Dbw} pada metode KM nilainya lebih kecil, yang berarti data ini cocok digunakan pada metode KM. Sedangkan pada data non linier indeks S_{Dbw} pada

metode KKM lebih kecil untuk semua data. Hal ini menunjukkan metode KKM lebih cocok digunakan pada jenis data non linier. Untuk analisis selanjutnya kernel yang digunakan adalah kernel Gaussian..

4.3 Penerapan metode *K-Means* dan *Kernel K-Means* untuk Pengelompokan Kabupaten/Kota di Indonesia

Kerangka konseptual terkait dengan faktor-faktor risiko yang menyebabkan penyakit hipertensi dalam penelitian ini adalah individu. Sementara dalam penelitian ini pengelompokan didasarkan pada faktor-faktor yang berpengaruh pada kejadian hipertensi secara agregat per kabupaten/kota. Oleh karena itu dilakukan pemodelan untuk menunjukkan apakah kerangka konseptual individu itu juga berlaku pada data yang bersifat agregat. Untuk memodelkan faktor-faktor yang berpengaruh terhadap kejadian hipertensi digunakan analisis regresi linier berganda, dengan model penelitian yang akan diuji adalah sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{28} X_{28}$$

Y = Persentase penderita hipertensi menurut kabupaten/kota

X_1, X_2, \dots, X_{28} = adalah variabel penelitian sebagaimana pada Tabel 3.1.

Ketepatan fungsi regresi secara statistik setidaknya dapat diukur dari nilai koefisien determinasi, nilai statistik uji F, dan nilai statistik uji t. Diperoleh hasil analisis secara agregat sebagai berikut :

Tabel 4.3 Nilai Koefisien Determinasi (R^2)

R	R^2	$R^2_{Adjusted}$	Standar Error
0.719	0.518	0.489	2.617

Sumber : Hasil olah data

Tabel 4.4 Ringkasan Tabel ANOVA (Uji F)

<i>Model</i>	<i>Sum Square</i>	<i>df</i>	<i>Means Square</i>	<i>F</i>	<i>Sig</i>
Regression	3440.584	28	122.878	17.936	0.000
Residual	3206.260	468	6.851		
Total	6646.844	496			

Sumber : Hasil Olah Data

Tabel 4.5 Ringkasan Pengujian Regresi Secara Parsial (Uji t)

<i>Model</i>	<i>Unstandardized Coefficients</i>		<i>Standardized Coefficient</i>	<i>t</i>	<i>Sig</i>
	<i>B</i>	<i>Std. Error</i>	<i>Beta</i>		
(Constant)	0.377	1.808		0.209	0.835
X1	0.055	0.018	0.208	3.052	0.002
X2	-0.113	0.054	-0.187	-2.100	0.036
X3	0.127	0.047	0.270	2.713	0.007
X4	0.153	0.117	0.163	1.315	0.189
X5	0.032	0.118	0.048	0.270	0.787
X6	0.060	0.091	0.061	0.664	0.507
X7	0.057	0.083	0.065	0.692	0.489
X8	0.178	0.146	0.163	1.219	0.223
X9	0.074	0.210	0.051	0.351	0.726
X10	-0.455	0.160	-0.272	-2.846	0.005
X11	-0.117	0.200	-0.074	-0.583	0.560
X12	-0.046	0.125	-0.044	-0.371	0.710
X13	-0.005	0.184	-0.003	-0.027	0.978
X14	-0.220	0.298	-0.083	-0.737	0.462
X15	-0.064	0.269	-0.031	-0.238	0.812
X16	0.819	0.310	0.340	2.639	0.009
X17	-0.688	0.366	-0.144	-1.881	0.061
X18	-0.175	0.087	-0.189	-2.016	0.044
X20	0.062	0.059	0.083	1.039	0.299
X21	-0.015	0.040	-0.031	-0.379	0.705
X22	0.081	0.023	0.128	3.517	0.000
X23	-0.041	0.010	-0.159	-3.983	0.000
X24	-0.009	0.018	-0.017	-0.470	0.639
X25	-0.014	0.018	-0.029	-0.754	0.451
X26	0.061	0.013	0.181	4.604	0.000
X27	0.025	0.013	0.070	1.876	0.061
X28	0.077	.058	0.045	1.335	0.183

Sumber : Hasil Olah Data

Hasil uji F menunjukkan secara simultan variabel faktor-faktor risiko mempunyai pengaruh yang signifikan terhadap kejadian penyakit hipertensi. Sedangkan hasil uji parsial (uji t) menunjukkan meskipun tidak semua variabel signifikan, ada 10 variabel yang signifikan, yaitu X1, X2, X3, X10, X16, X18 yang termasuk dalam variabel kejiwaan, serta variabel X21, X22, X25, dan X28. Namun demikian, berdasarkan penelitian sebelumnya, seluruh variabel dalam

penelitian ini tetap dimasukkan ke dalam model yang selanjutnya akan digunakan dalam analisis kelompok.

Tujuan dari penelitian ini adalah mengelompokkan kabupaten/kota di Indonesia berdasarkan penduduk dengan faktor-faktor risiko penyebab penyakit hipertensi. Studi kasus dalam penelitian ini menggunakan dataset RISKESDAS 2013 se-Indonesia (497 kabupaten dan 28 variabel).

Tabel 4.6 Jumlah Kabupaten menurut Provinsi di Indonesia tahun 2013

Kode	Propinsi	Jmlh Kab/kota	Kode	Propinsi	Jmlh Kab/kota
11	Nanggroe Aceh	23	52	Nusa Tenggara Barat	10
12	Sumatera Utara	33	53	Nusa Tenggara Timur	21
13	Sumatera Barat	19	61	Kalimantan Barat	14
14	Riau	12	62	Kalimantan Tengah	14
15	Jambi	11	63	Kalimantan Selatan	13
16	Sumatera Selatan	15	64	Kalimantan Timur	14
17	Bengkulu	10	71	Sulawesi Utara	15
18	Lampung	14	72	Sulawesi Tengah	11
19	Kep. Babel	7	73	Sulawesi Selatan	24
21	Kep. Riau	7	74	Sulawesi Tenggara	12
31	DKI. Jakarta	6	75	Gorontalo	6
32	Jawa Barat	26	76	Sulawesi Barat	5
33	Jawa Tengah	35	81	Maluku	11
34	D.I. Yogyakarta	5	82	Maluku Utara	9
35	Jawa Timur	38	91	Papua Barat	11
36	Banten	8	94	Papua	29
51	Bali	9	Total		497

Sumber : BPS RI

Tabel 4.7 merupakan deskripsi dari variabel-variabel yang digunakan dalam penelitian ini. Berdasarkan tabel tersebut dapat dilihat bahwa dari variabel faktor kejiwaan (X1-X20), variabel yang paling banyak dikeluhkan adalah variabel seringnya menderita sakit kepala (X1), dengan rata-rata penduduk yang mengalami keluhan sebanyak 29,1691 persen.

Tabel 4.7 Statistik Deskriptif Variabel Penelitian

Variabel	Minimum	Maximum	Mean	Std. Deviation
X1	0.99	75.19	29.169	13.885
X2	0.16	37.80	9.683	6.040
X3	0.31	52.69	13.691	7.801
X4	0.00	29.52	4.386	3.883
X5	0.00	39.95	6.251	5.505
X6	0.16	35.73	4.500	3.693
X7	0.00	37.63	4.568	4.137
X8	0.00	34.74	3.219	3.350
X9	0.00	31.35	2.383	2.540
X10	0.00	18.44	2.177	2.191
X11	0.00	27.46	2.465	2.334
X12	0.00	29.86	3.183	3.445
X13	0.00	29.94	2.264	2.106
X14	0.00	19.11	1.349	1.381
X15	0.00	21.09	1.590	1.747
X16	0.00	15.96	1.368	1.521
X17	0.00	10.42	0.707	0.766
X18	0.00	35.40	3.814	3.943
X19	0.00	40.12	5.753	4.959
X20	0.16	47.39	10.594	7.523
X21	8.20	54.09	35.321	5.752
X22	9.93	92.41	41.267	14.330
X23	43.73	99.10	87.042	7.184
X24	0.00	56.58	7.223	7.569
X25	0.05	61.99	11.992	10.765
X26	0.41	62.98	15.887	10.307
X27	0.00	23.06	2.006	2.113
X28	5.64	53.47	25.859	7.210

Sumber : Hasil Olah Data

Variabel yang mempengaruhi faktor kejiwaan tertinggi berikutnya adalah variabel sulit tidur dan mudah lelah, dengan rata-rata persentase penduduk yang mengalami keluhan tersebut masing-masing sebesar 13.6914 persen dan 10.5947 persen.

Faktor gaya hidup, dalam hal ini rata-rata penduduk yang merokok (X21) persentasenya mencapai 35.3214 persen. Sedangkan variabel yang berhubungan dengan aktivitas fisik, rata-rata penduduk yang melakukan aktivitas berat dan sedang masing-masing mencapai 41.2672 persen dan 87.0427 persen. Variabel yang berhubungan dengan pola makan, untuk variabel persentase penduduk yang mengkonsumsi makanan berlemak/berkolesterol/gorengan, dengan rata-rata persentase sebesar 11.9923 persen, dan untuk persentase penduduk yang mengkonsumsi kopi mencapai 15.8877 persen. Sedangkan persentase penduduk yang mengkonsumsi makanan asin hanya sebesar 7.2232 persen.

Variabel yang berhubungan dengan faktor lainnya yang mempengaruhi risiko penyakit hipertensi adalah variabel persentase penduduk yang mengalami obesitas. Persentase penduduk yang mengalami obesitas sebesar 25.8598 persen.

4.3.1 Rekomendasi Jumlah Kelompok (*Cluster*) yang Optimal

Untuk menentukan jumlah kelompok yang optimal, perlu melakukan perbandingan dari beberapa indeks validitas kelompok untuk mengevaluasi hasil analisis kelompoknya secara kuantitatif sehingga didapatkan jumlah kelompok yang optimum. Indeks validitas kelompok yang akan digunakan untuk menentukan jumlah kelompok dalam penelitian ini menggunakan indeks validitas *Dunn*, *Davies-Bouldin*, *Calinski-Harabasz*, *Xie-Beni*, dan *S_Dbw*.

Tabel 4.8 Indeks Validitas Kelompok untuk Menentukan jumlah Kelompok Optimal

Jumlah Kelompok	CH	D	DB	S_Dbw	XB
2	113.341	0.069	1.932	1.02114	6.805429
3	75.698	0.072	1.781	NA	5.85421
4	63.868	0.051	2.602	NA	11.2078
5	56.271	0.057	2.191	NA	8.657071
6	46.529	0.057	2.014	NA	8.399768
7	38.216	0.056	3.087	NA	8.595347
8	39.536	0.051	2.241	NA	9.938776
9	35.428	0.057	2.724	NA	7.831397
10	34.054	0.057	2.312	NA	7.598077

Sumber : Hasil Olah Data

Dari beberapa indeks di atas, jumlah kelompok yang optimal adalah sebanyak 2 dan 3 kelompok, ditunjukkan dengan nilai dari tiap-tiap indeks. Indeks CH dan Dunn (D), dengan nilai terbesar menunjukkan kelas yang optimal. Sedangkan indeks DB, S_Dbw, dan XB dengan nilai terkecil menunjukkan kelas yang optimal . Namun menurut Abonyi dan Feil dalam Syoer (2011), untuk menentukan berapa jumlah kelompok yang dipilih dari beberapa indeks tersebut, dalam beberapa literatur disebutkan: pertama, tidak ada indeks yang terbaik, sehingga referensi dari beberapa indeks validitas bisa dijadikan “acuan” untuk penentuan jumlah kelompok. Kedua, sesuai dengan keperluan penelitian, atau sesuai dengan kasus yang akan digunakan untuk tujuan peneliti (Syoer, 2011). Berdasarkan dua hal tersebut dan berdasarkan Tabel 4. 8 , maka dalam pembahasan berikutnya akan dibandingkan pengelompokan dengan metode KM dan KKM dengan jumlah kelompok sebanyak dua dan tiga kelompok.

4.3.2 Pengelompokan Kabupaten/Kota di Indonesia menggunakan metode KM dan KKM

a. Pengelompokan dengan metode KM

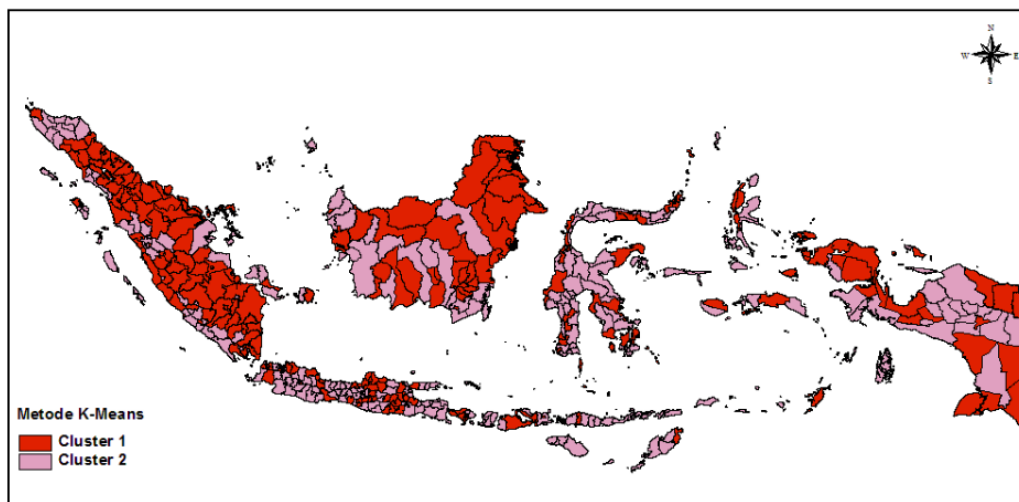
Hasil pengelompokan dengan $k=2$ terhadap 497 kabupaten/kota di Indonesia, diperoleh dua kelompok dengan ukuran 319 dan 178, dengan keanggotaan dapat dilihat pada Lampiran 2a. Nilai pusat masing-masing kelompok seperti pada tabel berikut :

Tabel. 4.9 Nilai Pusat Kelompok hasil pengelompokan dengan KM dengan k=2

Variabel	Kelompok 1	Kelompok 2
(1)	(2)	(3)
X1	21.467	42.973
X2	6.601	15.206
X3	9.573	21.073
X4	2.622	7.549
X5	3.688	10.845
X6	2.761	7.619
X7	2.777	7.779
X8	1.904	5.576
X9	1.447	4.062
X10	1.296	3.757
X11	1.565	4.078
X12	1.901	5.483
X13	1.475	3.681
X14	0.879	2.192

Variabel	Kelompok 1	Kelompok 2
(1)	(2)	(3)
X15	0.962	2.717
X16	0.854	2.291
X17	0.483	1.108
X18	2.082	6.918
X19	3.420	9.934
X20	6.750	17.485
X21	34.072	37.560
X22	39.708	44.062
X23	86.653	87.741
X24	6.841	7.909
X25	10.970	13.824
X26	14.158	18.988
X27	1.908	2.183
X28	25.876	25.831

Sumber : Hasil Olah Data



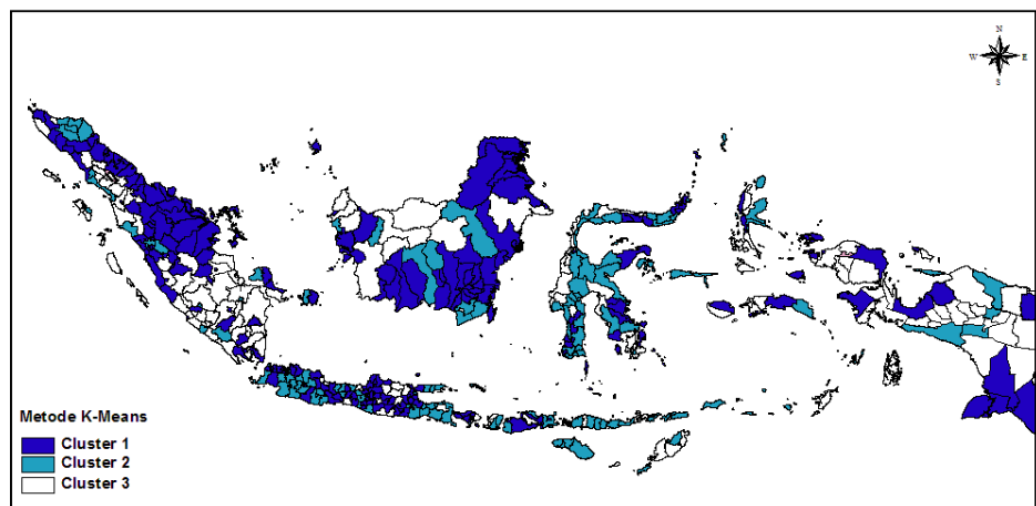
Gambar 4.3 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan K-Means Clustering dengan jumlah Kelompok sebanyak 2 Kelompok

Tabel. 4.10 Nilai Pusat Kelompok hasil pengelompokan dengan KM dengan k=3

Variabel	Kelompok 1	Kelompok 2	Kelompok 3
(1)	(2)	(3)	(4)
X1	23.740	45.083	22.105
X2	7.586	16.446	6.285
X3	11.217	22.572	8.700
X4	3.292	8.534	1.940
X5	4.762	12.177	2.617
X6	3.319	8.527	2.353
X7	3.654	8.618	1.889
X8	2.413	6.298	1.391
X9	1.849	4.516	1.072
X10	1.670	4.237	0.895
X11	1.784	4.560	1.473
X12	2.364	6.202	1.448
X13	1.654	4.128	1.392
X14	1.014	2.481	0.750

Variabel	Kelompok 1	Kelompok 2	Kelompok 3
(1)	(2)	(3)	(4)
X15	1.182	3.063	0.758
X16	1.046	2.560	0.681
X17	0.564	1.229	0.410
X18	2.667	7.881	1.555
X19	4.377	10.842	2.814
X20	8.176	19.048	5.964
X21	34.509	38.378	33.529
X22	32.039	43.959	55.761
X23	85.830	87.638	88.686
X24	6.614	7.959	7.573
X25	13.401	13.974	7.176
X26	12.769	19.492	17.854
X27	1.989	2.252	1.773
X28	28.212	25.805	21.478

Sumber : Hasil Olah Data



Gambar 4.4 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan K-Means Clustering dengan jumlah Kelompok sebanyak 3 Kelompok

Sedangkan pengelompokan dengan $k=3$, diperoleh kelompok dengan ukuran 236, 136, dan 125, dengan keanggotaan terlampir pada lampiran.2b. Nilai pusat kelompok tersaji pada tabel 4.10.

b. Pengelompokan dengan metode KKM

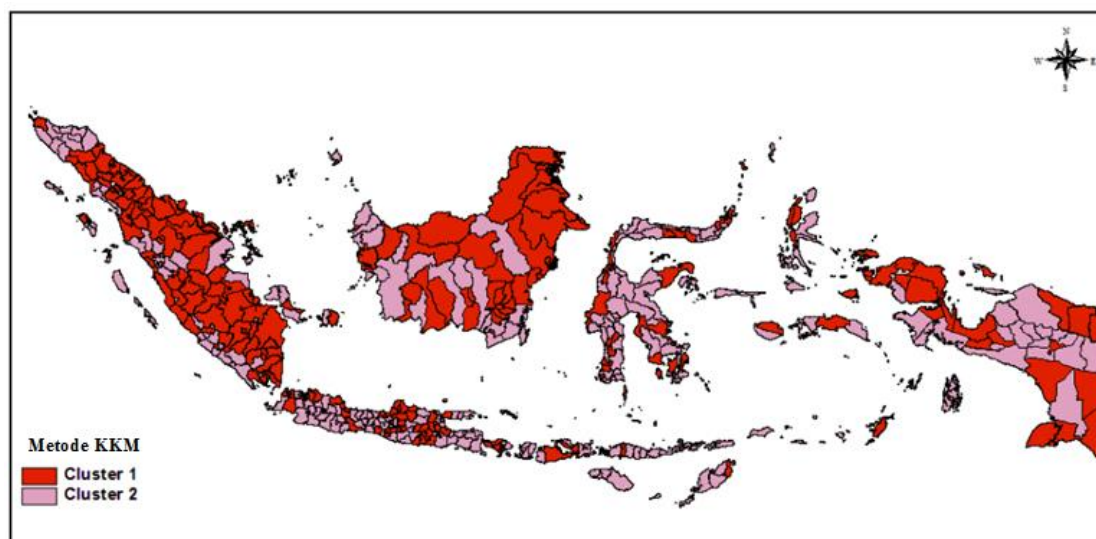
Hasil pengelompokan dengan $k=2$ diperoleh kelompok dengan ukuran 254 dan 243, dengan keanggotaan terlampir pada lampiran. 2c. Sedangkan nilai pusat kelompok dapat dilihat pada tabel berikut :

Tabel. 4.11 Nilai Pusat Kelompok hasil pengelompokan dengan KKM dengan $k=2$

Variabel	Kelompok 1	Kelompok 2
(1)	(2)	(3)
X1	20.110	38.638
X2	6.285	13.235
X3	9.286	18.296
X4	2.556	6.300
X5	3.672	8.947
X6	2.636	6.450
X7	2.761	6.457
X8	1.860	4.640
X9	1.424	3.386
X10	1.300	3.095
X11	1.508	3.466
X12	1.887	4.540
X13	1.411	3.158
X14	0.862	1.858

Variabel	Kelompok 1	Kelompok 2
(1)	(2)	(3)
X15	0.948	2.261
X16	0.824	1.937
X17	0.476	0.948
X18	2.046	5.662
X19	3.275	8.344
X20	6.494	14.881
X21	34.176	36.519
X22	36.746	45.993
X23	86.617	87.487
X24	6.051	8.449
X25	10.120	13.950
X26	12.447	19.484
X27	1.791	2.233
X28	26.620	25.065

Sumber : Hasil Olah Data



Gambar 4.5 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan Kernel K-Means *Clustering* dengan jumlah Kelompok sebanyak 2 Kelompok

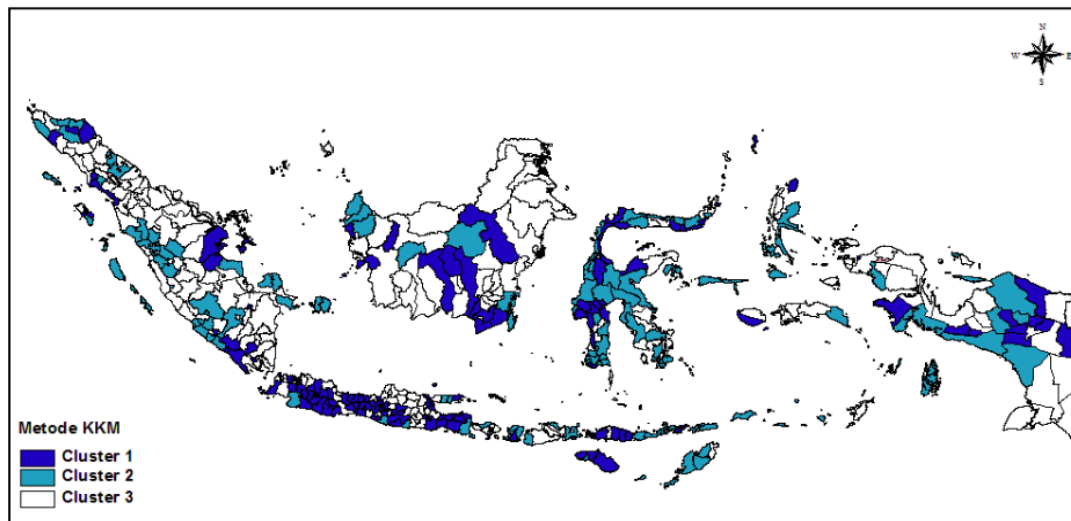
Sedangkan pengelompokan dengan $k=3$, diperoleh kelompok dengan ukuran 148, 112, dan 237, dengan keanggotaan terlampir pada lampiran 2d. Nilai pusat kelompok tersaji pada tabel berikut :

Tabel. 4.12 Nilai Pusat Kelompok hasil pengelompokan dengan KKM $k=3$

Variabel	Kelompok 1	Kelompok 2	Kelompok 3
(1)	(2)	(3)	(4)
X1	38.012	37.795	19.571
X2	14.352	11.253	6.026
X3	19.926	15.738	8.831
X4	7.353	4.707	2.382
X5	10.924	6.153	3.380
X6	7.424	4.810	2.529
X7	8.124	4.095	2.571
X8	5.918	2.803	1.731
X9	4.221	2.185	1.330
X10	3.795	2.084	1.212
X11	4.058	2.495	1.457
X12	5.930	2.622	1.734
X13	3.600	2.396	1.369
X14	2.240	1.269	0.830

Variabel	Kelompok 1	Kelompok 2	Kelompok 3
(1)	(2)	(3)	(4)
X15	2.863	1.402	0.884
X16	2.435	1.235	0.765
X17	1.149	0.667	0.451
X18	6.977	3.668	1.908
X19	9.924	5.675	3.185
X20	17.093	11.177	6.262
X21	36.629	36.236	34.073
X22	41.093	50.833	36.855
X23	87.487	87.911	86.355
X24	10.988	4.700	6.065
X25	19.486	7.150	9.601
X26	19.103	17.849	12.953
X27	2.217	2.085	1.839
X28	27.045	23.266	26.345

Sumber : Hasil Olah Data



Gambar 4.6 Peta Kabupaten/Kota di Indonesia Hasil Pengelompokan dengan Kernel K-Means *Clustering* dengan jumlah Kelompok sebanyak 3 Kelompok

Jika dibandingkan hasil pengelompokan antara metode KM dan KKM, pada pengelompokan dengan jumlah kelompok 2, perbedaan keanggotaan masing-masing kelompok pada dua metode ini terdapat pada 63 kabupaten/kota, dimana pada pengelompokan dengan metode KM dimasukkan dalam kelompok 1, pada KKM dimasukkan dalam kelompok 2. Kelompok 2 merupakan kelompok dengan persentase penduduk yang memiliki faktor risiko penyakit hipertensi yang lebih tinggi daripada kelompok 1, yaitu kelompok dengan persentase penduduk yang memiliki faktor risiko penyebab penyakit hipertensi lebih tinggi dari rata-rata.

Sedangkan pada pengelompokan dengan $k=3$, perbedaan keanggotaan jauh lebih bervariasi. Kelompok 1 merupakan kelompok dengan persentase penduduk yang memiliki faktor risiko penyakit hipertensi yang paling tinggi. Sedangkan kelompok 2 memiliki persentase penduduk dengan faktor risiko penyebab penyakit hipertensi dengan persentase di bawah kelompok 1, namun masih di atas rata-rata, sedangkan kelompok 3 merupakan kelompok dengan persentase penduduk yang memiliki faktor risiko penyebab penyakit hipertensi yang lebih rendah dari rata-rata.

4.3.3 Interpretasi Hasil Pengelompokan

Berdasarkan pembahasan sebelumnya dan berdasarkan pengujian ukuran validitas kelompok terhadap data penelitian juga menunjukkan penelompokan dengan metode KKM lebih baik dibandingkan metode KM. Maka pada bagian ini akan dilakukan interpretasi hasil pengelompokan kabupaten/kota di Indonesia berdasarkan penduduk dengan faktor risiko penyakit hipertensi menggunakan metode KKM.

Tabel. 4.13 Jumlah Anggota per Kelompok pada pengelompokan dengan metode KKM dengan $k=2$

Kelompok	Jumlah Anggota (Kab/kota)	%
Kelompok 1	254	51.11
Kelompok 2	243	48.89

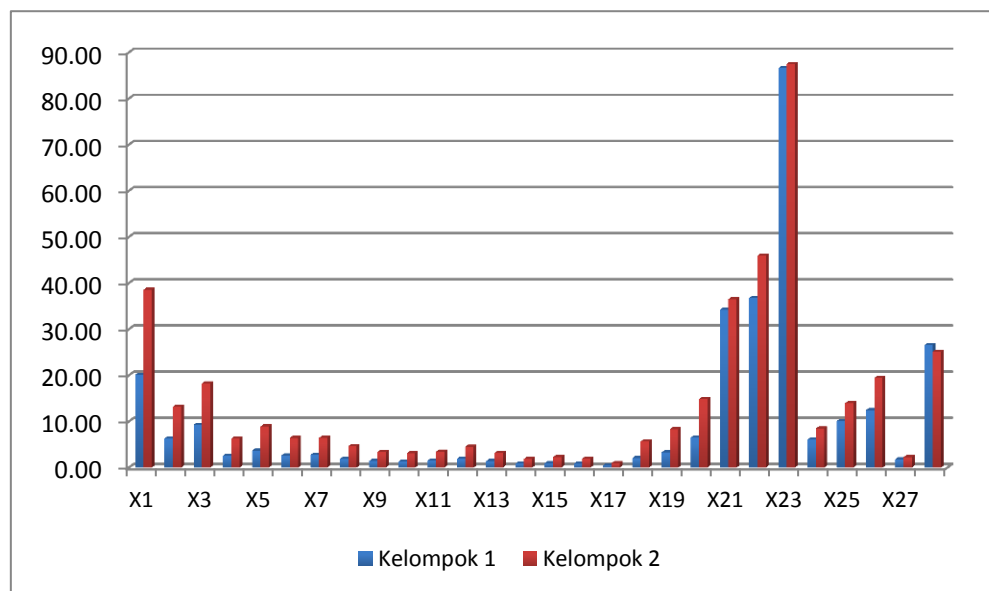
Sumber : Hasil Olah Data

Dari hasil pengelompokan menggunakan metode KKM dengan jumlah kelompok sebanyak 2, dengan keanggotaan masing-masing pada lampiran 2c, diperoleh urutan pengelompokan sebagai berikut :

- a. kelompok 2 sebanyak 243 kabupaten/kota (48.89 persen). Kelompok kedua merupakan kabupaten/kota di Indonesia dengan penduduk yang memiliki faktor risiko penyakit hipertensi yang lebih tinggi dilihat dari persentase penduduk dengan faktor risiko penyakit hipertensi menurut kabupaten/kota, antara lain persentase penduduk yang sering menderita sakit kepala pada kelompok 2 lebih tinggi dibandingkan pada kelompok 1. Hampir semua variabel, rata-rata persentase penduduk dengan faktor risiko hipertensi pada kelompok kedua lebih tinggi dibandingkan pada kelompok pertama, kecuali pada penduduk yang menderita obesitas (X_{28}), dimana pada kelompok pertama rata-rata persentasenya lebih tinggi persentasenya. Wilayah dengan persentase penduduk dengan faktor risiko tertinggi pada kelompok ini adalah Propinsi Nusa Tenggara Timur, dimana 18 dari 21 kabupaten/kota masuk dalam kelompok 2 (85.71 persen).

- b. kelompok 1 : terdiri dari 254 kabupaten/kota (51.11 persen). Kelompok ini terdiri dari kabupaten/kota di Indonesia dengan penduduk yang memiliki risiko penyakit hipertensi lebih rendah dibandingkan kelompok yang lainnya.

Gambaran rata-rata persentase penduduk menurut faktor risiko penyebab penyakit hipertensi berdasarkan kelompok yang terbentuk disajikan dalam Gambar 4.7.



Gambar 4.7 Rata-rata Persentase Penduduk dengan Faktor Risiko Penyebab Penyakit Hipertensi menurut Kelompok dengan $k=2$

Sedangkan hasil pengelompokan menggunakan metode KKM dengan jumlah kelompok sebanyak 3, dengan keanggotaan masing-masing pada lampiran 2d, diperoleh urutan pengelompokan sebagai berikut :

- a. kelompok 1 : terdiri dari 148 kabupaten/kota (29.78 persen). Pada kelompok 1, hampir semua variabel memiliki rata-rata persentase penduduk dengan faktor risiko yang lebih tinggi dibandingkan kelompok 2 dan 3, kecuali pada variabel X_{22} (intensitas aktivitas berat) dan X_{23} (intensitas aktivitas sedang), dimana kedua variabel ini lebih tinggi pada kelompok 2. Jika pengelompokan dibagi menjadi

kelompok dengan persentase tertinggi, tertinggi kedua, dan terendah, maka kelompok 1 merupakan kelompok yang memiliki persentase penduduk dengan faktor risiko penyebab penyakit hipertensi tertinggi.

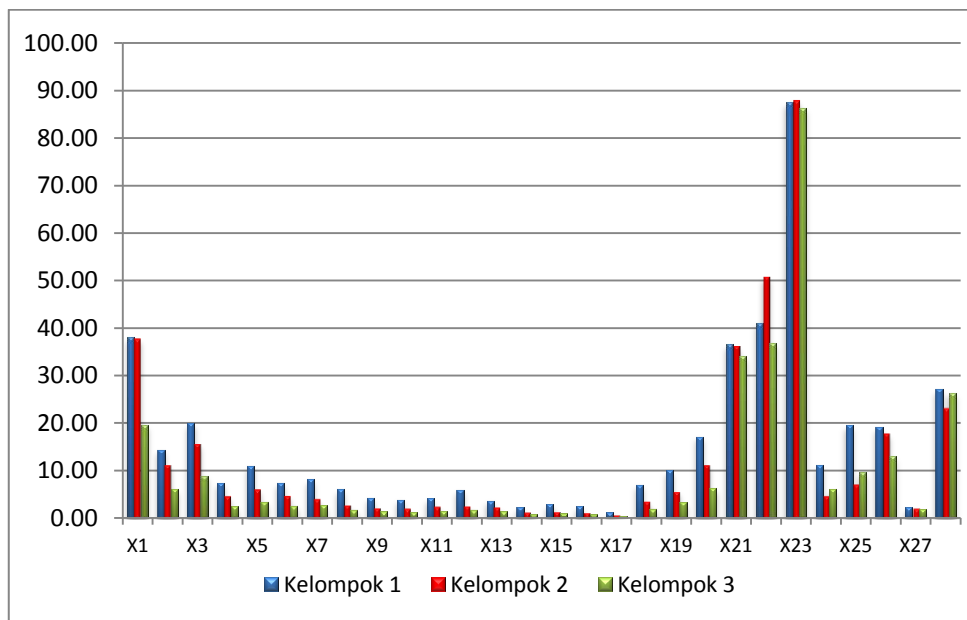
- b. kelompok 2 sebanyak 112 kabupaten/kota (22.54 persen). Kelompok kedua merupakan kabupaten/kota di Indonesia dengan penduduk yang memiliki faktor risiko penyakit hipertensi yang tertinggi kedua dilihat dari persentase penduduk dengan faktor risiko penyakit hipertensi menurut kabupaten/kota. Pada kelompok ini variabel X_{22} (intensitas aktivitas berat) dan X_{23} (intensitas aktivitas sedang) memiliki persentase tertinggi dibandingkan kelompok yang lain.
- c. Kelompok 3 terdiri dari 237 kabupaten/kota (47.69 persen), adalah kelompok kabupaten/kota dengan persentase penduduk dengan faktor risiko penyebab penyakit hipertensi terendah dibandingkan dengan kelompok yang lain.

Gambaran rata-rata persentase penduduk menurut faktor risiko penyebab penyakit hipertensi berdasarkan kelompok yang terbentuk disajikan dalam Gambar 4.8. Propinsi NTT, sebagian besar wilayah di Pulau Sulawesi dan Pulau Jawa mempunyai persentase penduduk dengan faktor risiko penyebab penyakit hipertensi yang cukup tinggi.

Tabel. 4.14 Jumlah Anggota per Kelompok pada pengelompokan dengan metode KKM dengan $k=3$

Kelompok	Jumlah Anggota (Kab/kota)	%
Kelompok 1	148	29.78
Kelompok 2	112	22.54
Kelompok 3	237	47.69

Sumber : Hasil Olah Data



Gambar 4.8 Rata-rata Persentase Penduduk dengan Faktor Risiko Penyebab Penyakit Hipertensi menurut Kelompok dengan k=3

Sebagaimana pembahasan di atas, pengelompokan kabupaten/kota dengan k=3 lebih menggambarkan variasi, sehingga pada bagian ini juga akan dibahas pengelompokan kabupaten/kota di Indonesia menurut dimensi faktor risiko penyebab penyakit hipertensi, yaitu menurut faktor kejiwaan, pola makan, aktivitas fisik, gaya hidup, dan lain-lain dengan k=3. Adapun hasil pengelompokan kabupaten/kota menurut dimensi faktor risiko tersebut adalah :

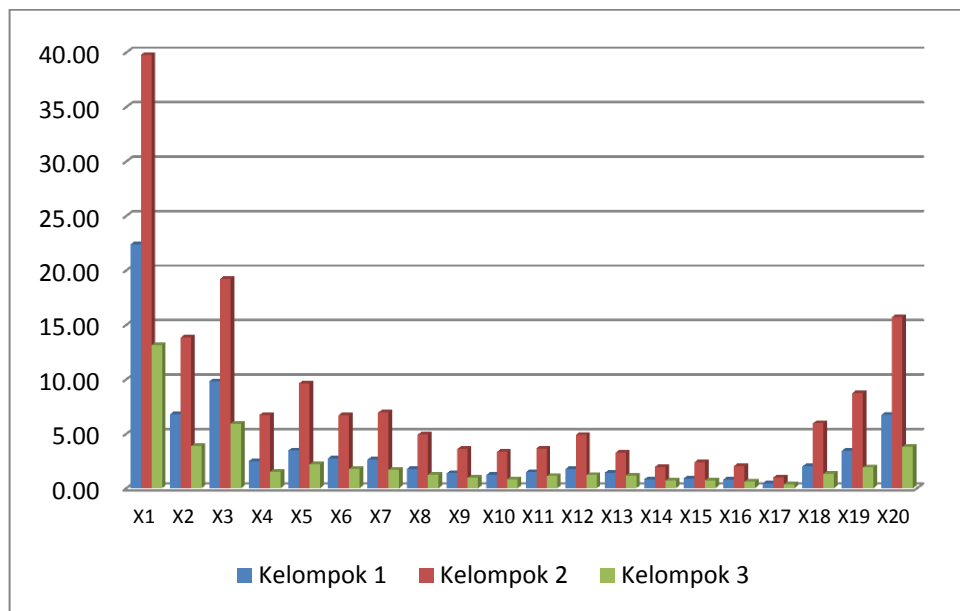
a. Pengelompokan kabupaten/kota di Indonesia menurut faktor kejiwaan

Berdasarkan hasil pengelompokan dengan metode KKM pada faktor kejiwaan, diperoleh hasil pengelompokan sebagai berikut :

Tabel. 4.15 Jumlah Anggota per Kelompok pada pengelompokan menurut faktor kejiwaan menggunakan metode KKM dengan k=3

Kelompok	Jumlah Anggota (Kab/kota)	%
Kelompok 1	159	31.99
Kelompok 2	244	49.09
Kelompok 3	94	18.91

Sumber : Hasil Olah Data



Gambar 4.9 Rata-rata Persentase Penduduk dengan Faktor Kejiwaan Penyebab Penyakit Hipertensi menurut Kelompok dengan $k=3$

Dengan melihat rata-rata persentase penduduk yang memiliki faktor risiko kejiwaan menurut kelompok sebagaimana Gambar 4.9, dapat diinterpretasikan karakteristik dari masing-masing kelompok sebagai berikut :

Kelompok 2 : Kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko kejiwaan tertinggi dibandingkan dengan kelompok yang lain. Pada kelompok ini, untuk seluruh faktor kejiwaan (X1-X20) memiliki rata-rata persentase lebih tinggi dari rata-rata pada variabel kejiwaan. Dapat diartikan kelompok ini terdiri dari kabupaten/kota dengan penduduk yang memiliki faktor risiko kejiwaan tertinggi yang dapat menyebabkan penyakit hipertensi.

Kelompok 1 : Pada kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko kejiwaan tertinggi kedua dibandingkan dengan kelompok yang lain. Pada kelompok ini, untuk seluruh faktor kejiwaan (X1-X20) memiliki rata-rata persentase tidak lebih tinggi dari rata-rata pada variabel kejiwaan, namun masih lebih tinggi jika

dibandingkan pada kelompok 3.. Dapat diartikan kelompok ini terdiri dari kabupaten/kota dengan penduduk yang memiliki faktor risiko kejiwaan tertinggi kedua yang dapat menyebabkan penyakit hipertensi. Jika pengelompokan dibagi menjadi kelompok tinggi, sedang, dan rendah, maka pada kelompok 1 merupakan kelompok kabupaten/kota dengan penduduk yang memiliki faktor risiko kejiwaan sedang yang dapat menyebabkan penyakit hipertensi.

Kelompok 3 : Sedangkan kelompok tiga adalah kelompok kabupaten/kota dengan persentase penduduk dengan faktor risiko kejiwaan terendah dibandingkan dengan kelompok yang lain.

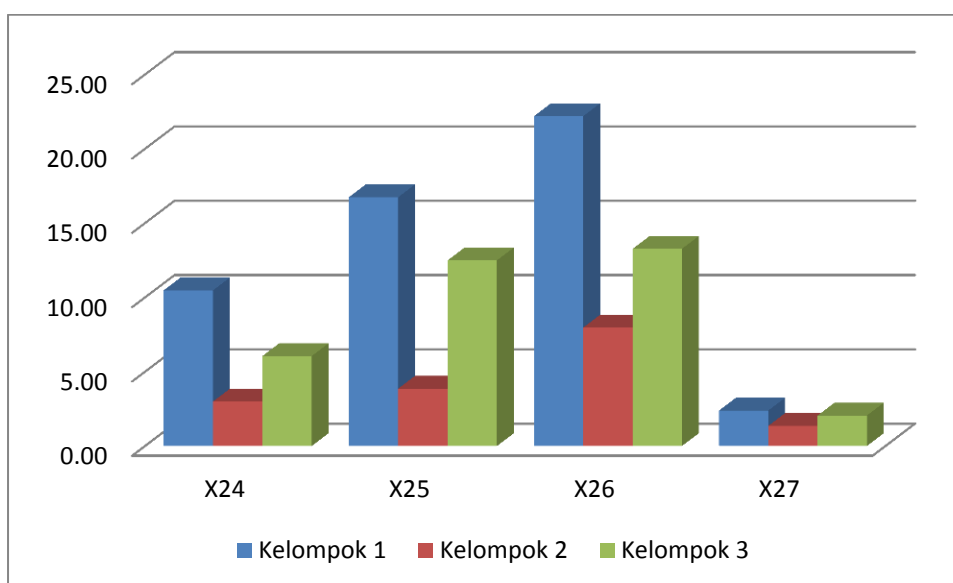
Keanggotaan masing-masing kelompok terlampir pada Lampiran 2e.

- b. Pengelompokan kabupaten/kota di Indonesia menurut faktor pola makan
Hasil pengelompokan dengan metode KKM pada faktor pola makan, diperoleh hasil pengelompokan sebagai berikut :

Tabel. 4.16 Jumlah Anggota per Kelompok pada pengelompokan menurut faktor pola makan menggunakan metode KKM dengan k=3

Kelompok	Jumlah Anggota (Kab/kota)	%
Kelompok 1	230	46.28
Kelompok 2	142	28.57
Kelompok 3	125	25.15

Sumber : Hasil Olah Data



Gambar 4.10 Rata-rata Persentase Penduduk dengan Faktor Pola Makan Penyebab Penyakit Hipertensi menurut Kelompok dengan $k=3$

Berdasarkan rata-rata persentase penduduk yang memiliki faktor risiko pola makan menurut kelompok (Gambar 4.10), dapat diinterpretasikan karakteristik dari masing-masing kelompok sebagai berikut :

Kelompok 1 : Kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko berupa faktor pola makan terburuk dibandingkan dengan kelompok yang lain. Pada kelompok ini, untuk seluruh faktor pola makan, yakni persentase penduduk yang tinggi konsumsi makanan asin, makanan berlemak, dan yang mengkonsumsi minuman berkafein (X24-X27) memiliki rata-rata persentase lebih tinggi dari rata-rata pada variabel pola makan. Jika pengelompokan dibagi menjadi kelompok dengan pola makan terburuk pertama, terburuk kedua, dan terburuk ketiga, maka pada kelompok 1 merupakan kelompok kabupaten/kota dengan penduduk yang memiliki faktor risiko pola makan terburuk pertama yang dapat menyebabkan penyakit hipertensi. Dapat diartikan kelompok ini terdiri dari kabupaten/kota dengan penduduk yang memiliki faktor risiko

berupa faktor pola makan terburuk yang dapat menyebabkan penyakit hipertensi.

Kelompok 3 : Pada kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko berupa faktor pola makan terburuk kedua dibandingkan dengan kelompok yang lain. Pada kelompok ini, untuk seluruh faktor pola makan (X24-X27) memiliki rata-rata persentase tertinggi kedua. Dapat diartikan kelompok ini terdiri dari kabupaten/kota dengan penduduk yang memiliki faktor risiko pola makan terburuk kedua yang dapat menyebabkan penyakit hipertensi.

Kelompok 2 : Sedangkan kelompok 2 adalah kelompok kabupaten/kota dengan persentase penduduk dengan faktor risiko pola makan terburuk ketiga (paling rendah persentase penduduk yang memiliki pola makan buruk) dibandingkan dengan kelompok yang lain.

Keanggotaan masing-masing kelompok terlampir pada Lampiran 2f.

- c. Pengelompokan kabupaten/kota di Indonesia menurut faktor aktivitas fisik
- Sedangkan hasil pengelompokan dengan metode KKM pada faktor aktivitas fisik dengan keanggotaan kelompok tersaji pada Lampiran 2g, diperoleh hasil pengelompokan sebagai berikut :

Tabel. 4.17 Jumlah Anggota per Kelompok pada pengelompokan menurut faktor aktivitas fisik menggunakan metode KKM dengan $k=3$

Kelompok	Jumlah Anggota (Kab/kota)	%
Kelompok 1	143	28.77
Kelompok 2	184	37.02
Kelompok 3	170	34.21

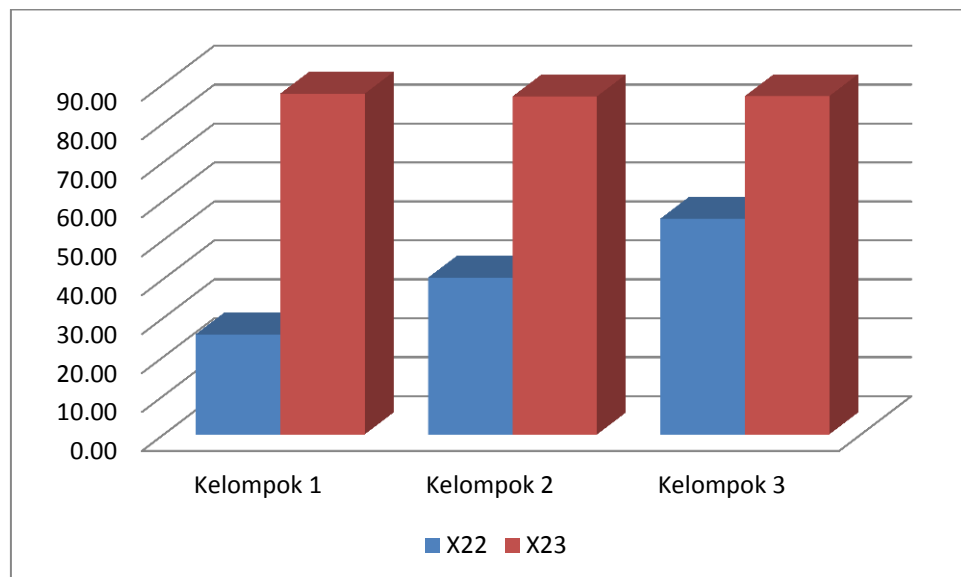
Sumber : Hasil Olah Data

Berdasarkan rata-rata persentase penduduk yang memiliki faktor risiko aktivitas fisik menurut kelompok (Gambar 4.11), dapat diinterpretasikan karakteristik dari masing-masing kelompok sebagai berikut :

Kelompok 3 : Kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko berupa faktor aktivitas fisik. Pada kelompok ini, untuk seluruh faktor aktivitas fisik, yakni persentase penduduk yang melakukan aktivitas fisik berat dan sedang (X22-X23) memiliki rata-rata persentase lebih tinggi pada variabel persentase penduduk yang melakukan aktivitas berat (X22) sedangkan persentase penduduk yang melakukan aktivitas sedang nilainya mendekati nilai rata-rata untuk semua kelompok.

Kelompok 2 : Pada kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko berupa faktor aktivitas fisik berat tertinggi kedua dibandingkan dengan kelompok yang lain.

Kelompok 1 : Sedangkan kelompok 2 adalah kelompok kabupaten/kota dengan persentase penduduk dengan faktor risiko aktivitas fisik berat tertinggi ketiga (paling rendah persentase penduduk yang melakukan aktivitas berat) dibandingkan dengan kelompok yang lain.



Gambar 4.11 Rata-rata Persentase Penduduk dengan Faktor Aktivitas Fisik Penyebab Penyakit Hipertensi menurut Kelompok dengan k=3

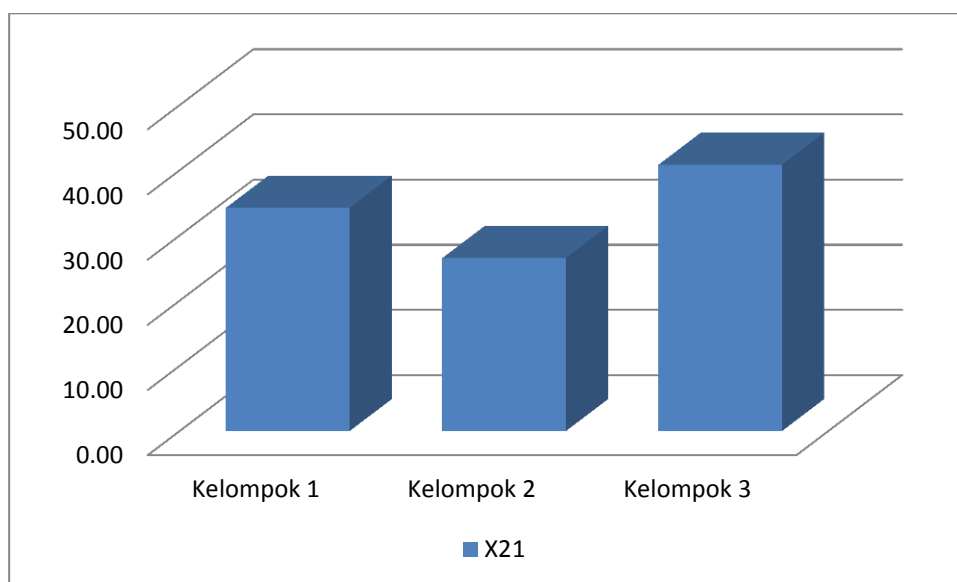
d. Pengelompokan kabupaten/kota di Indonesia menurut faktor gaya hidup

Metode KKM digunakan untuk data yang berdimensi tinggi. Maka pengelompokan dengan yang melibatkan 1 (satu) variabel tidak dapat menggunakan metode KKM, sebaliknya yang yang dapat digunakan adalah metode K-Means biasa. Maka pada dimensi faktor gaya hidup dan faktor lain-lain akan menggunakan metode K-Means biasa.

Tabel. 4.18 Jumlah Anggota per Kelompok pada pengelompokan menurut faktor gaya hidup menggunakan metode KKM dengan k=3

Kelompok	Jumlah Anggota (Kab/kota)	%
Kelompok 1	237	47.69
Kelompok 2	83	16.70
Kelompok 3	177	35.61

Sumber : Hasil Olah Data



Gambar 4.12 Rata-rata Persentase Penduduk dengan Faktor Gaya Hidup Penyebab Penyakit Hipertensi menurut Kelompok dengan k=3

Berdasarkan rata-rata persentase penduduk yang memiliki faktor risiko faktor gaya hidup merokok menurut kelompok sebagaimana Gambar

4.12, dapat diinterpretasikan karakteristik dari masing-masing kelompok sebagai berikut :

Kelompok 3 : Kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko berupa faktor gaya hidup merokok. Pada kelompok ini, memiliki rata-rata persentase penduduk yang merokok lebih tinggi daripada rata-rata variabel persentase penduduk yang merokok (X21). Dibandingkan kelompok yang lain kabupaten/kota pada kelompok ini memiliki persentase penduduk yang merokok paling tinggi.

Kelompok 1 : Pada kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk dengan faktor risiko berupa faktor gaya hidup merokok tertinggi kedua dibandingkan dengan kelompok yang lain.

Kelompok 2 : Sedangkan kelompok 2 adalah kelompok kabupaten/kota dengan persentase penduduk dengan faktor risiko gaya hidup merokok tertinggi ketiga (paling rendah persentase penduduk yang merokok) dibandingkan dengan kelompok yang lain.

Keanggotaan untuk masing-masing kelompok tersaji pada Lampiran 2h.

e. Pengelompokan kabupaten/kota di Indonesia menurut faktor lain-lain

Hasil pengelompokan kabupaten/kota di Indonesia menurut faktor lain-lain, dalam hal ini faktor penduduk yang mengalami obesitas tersaji dalam tabel 4.19 berikut ini :

Tabel. 4.19 Jumlah Anggota per Kelompok pada pengelompokan menurut faktor lain-lain menggunakan metode KKM dengan $k=3$

Kelompok	Jumlah Anggota (Kab/kota)	%
Kelompok 1	103	20.72
Kelompok 2	223	44.87
Kelompok 3	171	34.41

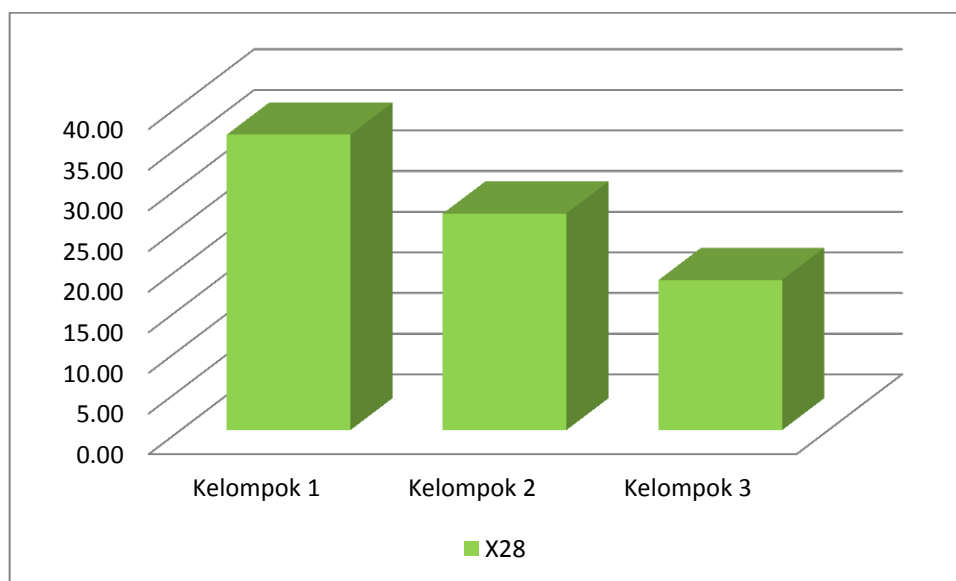
Sumber : Hasil Olah Data

Karakteristik hasil pengelompokan berdasarkan rata-rata persentase penduduk yang memiliki indeks massa tubuh lebih besar dari 25 (mengalami obesitas) adalah sebagai berikut :

Kelompok 1 : Kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk yang mengalami obesitas lebih tinggi daripada rata-rata variabel persentase penduduk yang mengalami obesitas (X28). Dibandingkan kelompok yang lain kabupaten/kota pada kelompok ini memiliki persentase penduduk yang mengalami obesitas paling tinggi.

Kelompok 2 : Pada kelompok ini terdiri dari kabupaten/kota dengan persentase penduduk yang mengalami obesitas tertinggi kedua dibandingkan dengan kelompok yang lain.

Kelompok 3 : Sedangkan kelompok 2 adalah kelompok kabupaten/kota dengan persentase penduduk dengan faktor lain-lain tertinggi ketiga (paling rendah persentase penduduk yang mengalami obesitas) dibandingkan dengan kelompok yang lain. Keanggotaan masing-masing kelompok tersaji pada Lampiran 2i.



Gambar 4.13 Rata-rata Persentase Penduduk dengan Faktor Lain-lain Penyebab Penyakit Hipertensi menurut Kelompok dengan k=3

BAB V

KESIMPULAN DAN SARAN

Berdasarkan uraian pada bab-bab sebelumnya dan hasil serta pembahasan, dapat dirumuskan kesimpulan dan saran sebagai berikut :

5.1 Kesimpulan

1. Algoritma Kernel K-Means *Clustering* pada prinsipnya mirip dengan metode K-Means *Clustering*, perbedaannya terletak pada perubahan masukan (inputnya), dimana dalam Kernel K-Means, data dipetakan pada dimensi baru yang lebih tinggi menggunakan fungsi kernel untuk selanjutnya dilakukan proses *clustering*.
2. Berdasarkan studi simulasi, untuk jumlah pengamatan besar (pada penelitian ini n sebanyak 3000), metode Kernel K-Means memberikan hasil yang lebih baik untuk data yang tidak dapat dipisahkan secara linier maupun yang dapat dipisahkan secara linier. Hal ini dapat dilihat dari hasil *plotting* data simulasi. Selain itu nilai *clustering error* pada pengelompokan dengan metode Kernel K-Means menghasilkan nilai yang lebih kecil dibandingkan dengan metode K-Means, terlebih untuk jumlah data yang lebih besar, dimana semakin besar jumlah amatan, akan memperkecil nilai *clustering error*. Berdasarkan indeks validitas kelompok, dengan indeks validitas S_{Dbw} , pada pengelompokan dengan Kernel K-Means menghasilkan nilai yang lebih kecil, sehingga metode ini dinilai lebih baik daripada K-Means. Dengan perbandingan beberapa kriteria pengelompokan tersebut di atas, pengelompokan menggunakan metode Kernel K-Means lebih baik dibandingkan dengan metode K-Means.
3. Pengelompokan kabupaten/kota menggunakan metode Kernel K-Means dengan jumlah kelompok sebanyak 3 (tiga) lebih menggambarkan variasi dari anggota kelompok. Hasil pengelompokan menggunakan metode Kernel K-Means pada faktor kejiwaan menunjukkan karakteristik kelompok dengan persentase penduduk dengan faktor risiko kejiwaan tertinggi berada pada kelompok. 2. Pengelompokan pada faktor pola makan

menunjukkan kelompok yang memiliki penduduk dengan faktor risiko pola makan terburuk berada pada kelompok 1. Kelompok kabupaten/kota yang memiliki penduduk dengan faktor risiko intensitas aktivitas fisik tertinggi berada pada kelompok 3. Sedangkan kelompok kabupaten/kota dengan persentase penduduk yang merokok dan mengalami obesitas masing-masing berada pada kelompok 3 dan 1.

5.2 Saran

Berdasarkan hasil penelitian ini, peneliti dapat memberikan saran sebagai berikut:

1. Data simulasi yang digunakan sebaiknya lebih dari satu populasi untuk membandingkan hasil pengelompokkannya, sehingga baik metode ataupun index validitasnya dapat dibandingkan dengan kasus lainnya .
2. Dalam penelitian ini, fungsi kernel yang digunakan berdasarkan konsep jarak *Euclidean*, dimana jarak Euclidean dapat digunakan ketika variabel diasumsikan memiliki varians yang sama atau tidak terdapat covarians antar variabel. Sehingga untuk penelitian selanjutnya untuk data yang memiliki perbedaan skala pengukuran dan terdapat korelasi pada variabelnya dapat menggunakan konsep jarak Mahalanobis.
3. Kebijakan untuk mengatasi atau mengurangi faktor risiko penyebab penyakit hipertensi ditekankan pada pengelompokan menurut faktor-faktor risiko di atas.

DAFTAR PUSTAKA

- Anggara, F., H, D, Prayitno, N. (2013). Faktor-faktor yang Berhubungan dengan Tekanan Darah di Puskesmas Telaga Murni, Cikarang Barat tahun 2012, *Jurnal Ilmiah Kesehatan* 5 (1).
- Arbelaitz, O., Gurrutxaga, I., Muquerza, J. Perez, J. M, & Perona, I. (2013), “An Extensive Comparatiave Syudy of Cluster Validity Indices”, *Pattern Recognition*, Vol. 46, hal. 243-256.
- Ariesen M.J, Claus SP., Rinkel G. J., Alqra A., (2003), “Risk Factor for intracerebral Hemorrhage in the General Population : a Systematic Review”, *Stroke*, 2003 : 34 hal. 2060-2065.
- Black, J.M., & Hawks, J.H. (2009). Medical Surgical Nursing : Clinical Management for Positives Outcomes. Elsevier : Philadelphia
- Buelens B., Pauly T., William R., Sale A., (2009), “Kernel Methods for the Detection and Classification of fish Schools in Single Beam and Multibeam Acoustic Data”, *ICES Journal of Marine Science*, 66 hal. 1130-1135.
- Cahyaningsih, R. (2008). *Analisis Pola Konsumsi Pangan di Provinsi Jawa Barat*, Skripsi, Institut Pertanian Bogor, Bogor.
- Dewi, A., (2012). *Metode Cluster Ensemble untuk Pengelompokan Desa Perdesaan di Propinsi Riau*, Tesis, Institut Teknologi Sepuluh Nopember (ITS), Surabaya.
- Dhillon, I. S., Guan Y., Kulis B.,(2005),
- Ferreira, M. R. P., & Carvalho, F. A. T. (2014), “ Kernel K-means Clustering Algorithm Based on an Adaptive Mahalanobis Kernel”, *International Joint Conference on Neural Networks (IJCNN)*, Beijing.
- Filippone, M., Camastra, F., Masuli, F., Rovetta, S. (2008). A Survey of Kernel and Spectral Methods for Clustering, *Pattern Recognition*, Vol. 41, hal 176-190 .
- Ginsberg, M. D., (2005), “Neuroprotection for Ischemic Stroke : Past, Present, and Future”, *Neuropharmacology*, 55 (3), hal 363-389.

- Girolami, M., (2002a), Mercer Kernel-Based Clustering in Feature Space, *Lecture handout : Machine Learning Module, University of Glasgow, Glasgow*.
- Girolami, M., (2002b), Mercer Kernel-Based Clustering in Feature Space. *Transaction on Neural Networks*, Vol.13, No 3, hal.780-784.
- Goldstein, et al., (2011). Guidelines for The Primary Prevention of Stroke, *Journal of the American Heart Association*.
- Handajani, A., Roosihermiatie, B., Maryani, H. (2012). Faktor-faktor yang Berhubungan dengan Pola Kematian pada Penyakit Degeneratif di Indonesia, *Buletin Penelitian Sistem Kesehatan*, Vol 13, No 1, hal 42-53.
- Hanifah, E., (2010), Metode Latent Class Clustering untuk Variabel Indikator Bertipe Campuran dalam Rangka Pengelompokan Desa, Tesis, Universitas Padjadjaran, Bandung.
- Hatma, R.J, Lukito, W, Rumawas, Y.S.P (2005), Fatty Acids Intake Among Diverse Ethnic Groups in Indonesia. *Med J Indonesia*, Vol 14, No.4 hal 242-248, Oktober-Desember 2005
- Jain, A. K, Dubes, R. C.,. (1988), *Algorithms for Clustering Data* . Michigan State University, Prentice Hall, New Jersey.
- Johnson, R. A., & Wichern, D. W., (2002), *Applied Multivariate Statistical Analysis*, (5th Ed), Prentice Hall, New Jersey.
- Kaplan, N. M., & Hyman, D., (1985), *Treatment of Patient with Mild Hypertension*, dalam *Hypertension*, American Heart Association, Dallas, hal 165-170.
- Kearney PM., Whelton M., Reynolds K., Myntber P., Whelton PK., He J.,(2005), Global Bueden of Hypertension : Analysis of Worldwide Data, *The Lancet*, Vol. 356, No. 9455, Hal. 217-223.
- Kissela B., et al (2004), Original Controbutions Srtoke in a Biracial Population The Excess Burden of Stroke Among Blacks, *Stroke*, 2004 ; 35, hal. 426-431.
- Kementerian Kesehatan RI. (2014a), *Profil Kesehatan Indonesia Tahun 2013*, Kemenkes RI, Jakarta.
- Kementerian Kesehatan RI, (2014b), *Laporan Riskesdas 2013*, Kemenkes RI, Jakarta.

- Kittner SJ, Stern BJ, Feeser BR, Hebel R, Nagey DA, Buchholz DW, Earley CJ, Johnson CJ, Macko RF, Sloan MA, Wityk RJ, Wozniak MA. (1996). *Pregnancy And The Risk Of Stroke*. N Engl J Med. 1996; 335: 768–774.
- Kovacs, F., Legany, C. & Babos, A. (2005), Cluster Validity Measurement Techniques. Hungaria.
- Liu, Y., Li, Z., Xiong, H., Gao, X., & WU, J. (2010), Understanding of Internal Clustering Validation Measures, *IEEE International Conference on Data Mining*, IEEE, New York.
- Marliani, L., Tantan S., (2007), *100 Question & Answer Hipertensi*, Elex Media Komputindo, Jakarta.
- Martha, A. (2012). *Analisis Faktor-faktor Risiko yang Berhubungan dengan Penyakit Diabetes Melitus pada Perusahaan X*, Tesis, Universitas Indonesia, Depok.
- Mattjik, A. A., & I Made S., (2011), *Sidik Peubah Ganda dengan Menggunakan SAS*, Kampus IPB Darmaga, Bogor.
- Mehar, A. M, Matawie, K. & Maeder, A. (2013), “Determining an Optimal Value of K in K-means Clustering”, *IEEE International Conference on Bioinformatics and Biomedicine*, IEEE, New York.
- Mudjiyanto, T.T, Hidayat, T.S, Hermina, dkk (1993), Pola Konsumsi Makanan Masyarakat Berpenghasilan Rendah di Wilayah Pengembangan Industri, *Penelitian Gizi Makan*, Vol. 16 Hal 59-69.
- Munaf, A. R. M. N. S. P, (2010), *Pengelompokan Kabupaten/Kota di Jawa Timur sebagai Dasar Perencanaan Pembiayaan Kegiatan Pendataan Lapangan Menggunakan Algoritma Fuzzy C-Means dengan Modifikasi Formula Jarak*, Tesis, ITS, Surabaya.
- Murfi, H. (2009), *Metode Kernel*, Bahan kuliah : Machine Learning, Universitas Indonesia, Depok.
- Nababan, D., (2008), *Hubungan Faktor Resiko dan Karakteristik Penderita dengan Kejadian Penyakit Jantung Koroner di RSUD. Pirngadi Medan Tahun 2008*, Tesis, Universitas Sumatera Utara.
- Ningrum, Nor Indah F., (2012), *Fuzzy C-Means Clustering dengan Analisis Robust*, Tesis, Universitas Padjadjaran, Bandung.

- Noor, .N. N, (2014), *Epidemiologi*, Rineka Cipta, Jakarta.
- Pleis JR, Lethbridge-Cejku M. (2006).” Summary Health Statistics For U.S. Adults. National Health Interview Survey”, *Vital Health Stat.* 10. 2007:1–153.
- Puspita WR., (2009), *Gaya Hidup pada Mahasiswa Penderita Hipertensi*, Skripsi, Fakultas Psikologi, Universitas Muhammadiyah Surakarta, Surakarta.
- Rochmayanti, (2011), *Analisis Faktor-Faktor yang Mempengaruhi Kualitas Hidup Pasien Penyakit Jantung Koroner di Rumah Sakit PELNI Jakarta*, Tesis, Universitas Indonesia.
- Rubattu S, Stanzione R, Gigante B, Bagalino A, Musumeci B, Volpe M. (2001). *Genetic Susceptibility To Cerebrovascular Accidents.* J Cardiovasc Pharmacol. 2001;38(suppl 2):S71–S74.
- Santosa, B., (2007), *Data Mining Terapan dengan Matlab*, Graha Ilmu, Yogyakarta.
- Saptiningsih, M. (2012). *Determinan Infeksi Saluran Kemih Pasien Diabetes Melitus Perempuan di RSB Bandung*, Tesis, Universitas Indonesia, Depok.
- Sharma S., (1996), *Applied Multivariate Techniques*, John Wiley & Sons, Inc.
- Sitorus, R. J., Hadisaputro, S., Kustiowati, E. (2008), *Faktor-Faktor yang Mempengaruhi Kejadian Stroke pada Usia Muda Kurang dari 40 Tahun (Studi Kasus Di Rumah Sakit Di Kota Semarang)*, *Epidemiologi*.
- Sugiharto, A. (2007), *Faktor-Faktro Risiko Hipertensi Grade II pada Masyarakat*, Tesis, Universitas Diponegoro, Semarang.
- Suryanto, E., (2015), *Perbandingan Reduced Support Vector Machine dan Smooth Support Vector Machine untuk Klasifikasi Large Data*, Skripsi, ITS, Surabaya.
- Sustrani, L.,(2006), *Hipertensi*, Jakarta : Gramedia.
- Syoer, R. R., (2011), *Analisis Kelompok dengan Algoritma Fuzzy C-Means dan Gath-Geva Clustering (Studi Kasus Pengelompokan Desa/Kelurahan di Kabupaten Kutai Kartanegara)*, Tesis, ITS, Surabaya.
- Touze E, Rothwell PM. (2008).” Sex Differences In Heritability Of Ischemic Stroke: A Systematic Review And Meta-Analysis”, *Stroke*. 2008;39:16–23.

- Turner ST, Boerwinkle E. (2003), “Genetics Of Blood Pressure, Hypertensive Complications, And Antihypertensive Drug Responses”, *Pharmacogenomics*, 2003;4:53–65.
- Tzortzis, G. F., Likas, A. C., (2009). “The Global Kernel k-Means Algorithm for Clustering in Feature Space”, *IEEE Transactions on Neural Networks*, Vol. 20, No.7, 1181-1194.
- Whiting, D. R., Guariguata, L., Weil, C., & Shaw, J. (2011). IDF Diabetes Atlas : Global Estimates of the Prevalence of Diabetes for 2011 and 2030. IDF.
- Wulandari, I., (2014), *Piecewise Polynomial Smooth Support Vector Machine untuk Klasifikasi Desa Tertinggal di Propinsi Kalimantan Timur tahun 2011*, Tesis, ITS, Surabaya.
- Xie, X. L, Beni, G., (1991), *A Validity Measure for Fuzzy Clustering*, IEEE PAMI, Vol. 13, No.8 , Hal 841-847.
- Zahuranec DB, Brown DL, Lisabeth LD, Gonzales NR, Longwell PJ, Eden SV, Smith MA, Garcia NM, Morgenstern LB. (2006), “Differences In Intracerebral Hemorrhage Between Mexican Americans And Non-Hispanic Whites”, *Neurology*, 2006;66:30–34.
- Zhao H., Liang J., Hu H., (2006), “Clustering Validity Based on Improved Hubert γ Statistic and the Separation of Clusters”, Proceeding of the First International Conference on Innovative Computing, Zinformation, and Control (ICICIC’06), IEEE Computer Society.

Lampiran 1. Data Penelitian

Kode Kab/ Kota	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25	X26	X27	X28
1101	32.58	9.25	11.32	2.77	3.28	2.51	1.73	2.07	1.90	0.78	1.90	1.38	1.82	1.04	0.86	1.30	0.43	1.64	2.59	3.63	33.79	57.04	91.44	2.07	6.40	28.18	1.99	21.61
1102	51.21	24.29	33.47	12.30	14.11	14.31	8.47	9.78	10.58	6.75	7.66	10.58	4.94	4.23	4.44	4.84	3.63	9.38	11.79	31.55	36.59	35.08	75.81	1.61	5.85	12.70	1.92	32.46
1103	28.67	12.31	14.10	4.14	5.64	2.82	1.69	1.88	1.32	1.32	1.88	1.69	2.91	1.50	1.41	1.22	1.03	3.48	4.23	7.61	31.86	32.24	75.75	4.70	2.73	21.80	0.94	25.75
1104	7.50	1.89	3.39	0.55	0.95	1.26	1.50	0.63	0.39	0.63	0.87	0.71	0.87	0.55	0.63	0.32	0.32	0.55	0.79	2.21	35.83	64.40	91.79	1.34	2.05	4.66	0.71	31.73
1105	55.57	23.74	28.09	19.66	22.98	11.40	20.77	11.15	8.51	9.19	6.72	8.68	5.70	3.06	3.32	5.11	2.21	13.45	21.87	28.34	37.87	44.51	86.64	6.47	12.17	14.55	1.45	28.85
1106	33.13	15.40	18.07	9.64	12.65	6.80	8.52	6.71	3.27	4.04	3.27	7.92	3.79	1.81	2.32	2.41	0.60	5.16	8.18	19.97	39.93	45.35	93.37	1.64	2.67	19.97	7.57	34.68
1107	30.79	9.50	10.89	4.55	6.26	2.76	2.44	3.09	2.68	2.52	2.68	2.27	2.92	1.62	1.30	1.87	0.65	3.17	6.01	10.48	34.93	41.19	88.55	7.23	24.45	27.86	2.68	24.86
1108	22.49	5.76	7.18	2.92	3.00	2.21	3.79	1.58	1.26	1.34	1.34	1.82	1.97	1.10	1.03	0.71	0.39	3.24	5.68	5.52	32.12	31.97	90.21	3.55	2.84	22.81	4.89	29.04
1109	29.43	11.48	16.25	9.22	9.14	6.14	9.78	4.61	3.64	4.85	2.75	3.15	2.75	1.21	1.37	1.13	0.49	6.55	10.43	15.44	29.67	33.47	76.07	3.40	5.82	21.75	2.67	29.18
1110	44.05	14.45	17.14	13.03	11.05	6.73	6.16	3.68	3.33	4.67	3.97	6.37	3.97	1.77	2.34	1.27	0.78	2.90	10.13	17.92	36.40	44.41	87.54	1.49	2.83	14.59	3.19	30.24
1111	39.45	10.02	14.87	7.01	8.17	6.24	9.01	2.62	1.85	1.77	2.23	3.00	2.54	1.23	0.69	0.54	0.23	6.16	11.86	22.88	31.82	51.62	84.05	8.78	6.09	22.96	7.94	27.50
1112	20.19	8.05	9.10	4.32	6.18	1.87	1.63	1.28	0.93	0.93	1.63	0.70	2.22	1.05	0.93	0.70	0.35	1.28	2.45	4.08	30.11	34.89	82.61	12.25	2.92	16.34	1.40	27.42
1113	24.30	8.07	8.07	2.97	3.53	3.62	1.67	1.39	0.74	1.30	1.58	1.95	1.30	0.56	0.65	0.46	0.28	1.58	2.04	4.36	38.96	36.83	90.17	1.67	8.53	25.70	0.37	27.46
1114	26.45	5.05	11.28	1.76	1.39	1.32	2.12	1.39	0.29	0.44	0.73	4.18	1.32	0.44	0.29	0.15	0.15	1.32	7.91	8.64	35.02	38.61	87.99	1.10	0.95	3.37	0.73	25.20
1115	26.66	12.18	12.18	3.58	3.07	2.98	3.92	1.70	1.19	1.02	1.53	1.19	1.36	0.77	0.77	0.85	0.60	2.64	3.75	12.78	27.17	27.26	54.43	9.11	2.56	17.38	5.28	17.97
1116	26.97	9.82	11.87	5.93	6.69	3.99	2.80	1.62	1.40	2.16	1.94	1.51	2.27	1.08	0.97	0.86	0.43	2.70	7.01	16.61	39.16	47.90	91.05	1.73	1.94	21.04	3.67	31.28
1117	45.01	27.40	32.72	17.52	23.14	12.88	17.04	14.91	10.16	9.00	9.00	16.36	8.81	4.36	8.33	6.58	1.36	14.71	18.30	27.20	41.53	53.34	93.80	3.58	5.52	32.72	1.36	33.40
1118	21.72	7.02	10.78	5.43	4.26	3.17	3.43	2.76	1.42	1.75	2.17	2.09	2.34	1.09	1.17	1.17	0.25	3.93	4.59	10.28	34.84	56.73	96.74	7.94	5.76	27.57	1.25	27.49
1171	17.86	5.75	10.33	2.17	2.87	1.24	3.26	1.79	1.16	0.85	1.63	2.10	1.71	1.01	0.78	0.93	0.31	0.93	3.11	8.77	32.30	25.39	96.04	1.32	4.89	11.10	1.63	34.47
1172	34.89	7.01	17.03	4.84	4.01	7.01	8.85	3.67	2.67	1.17	3.17	2.84	3.01	1.84	1.67	1.50	0.33	2.84	17.53	18.03	38.06	34.39	85.64	2.17	3.84	13.02	1.34	38.56
1173	9.29	4.57	5.57	2.71	3.33	2.01	2.32	1.70	1.78	3.10	2.01	1.63	1.47	1.47	1.39	1.24	0.70	2.17	2.55	3.95	33.51	22.68	86.38	1.24	1.78	6.19	1.01	32.28
1174	25.99	11.09	16.11	9.96	11.90	8.02	10.61	5.26	3.16	4.86	3.56	4.37	2.83	1.78	1.94	2.27	1.30	3.81	6.48	8.74	31.50	32.31	85.67	1.70	3.48	9.80	3.00	33.85
1175	58.31	29.74	32.55	12.88	19.44	17.21	14.05	11.48	8.90	5.74	8.20	8.43	7.38	3.04	6.21	5.62	2.11	8.43	12.06	32.44	38.17	45.67	77.99	18.50	14.40	22.48	6.44	29.16
1201	52.95	27.79	24.89	5.36	8.45	7.99	1.36	1.18	1.18	1.27	2.54	2.27	2.63	1.09	0.73	0.73	0.54	2.54	5.54	28.97	23.25	54.13	88.56	4.36	0.54	2.36	0.45	7.08
1202	20.29	9.18	8.82	1.94	1.51	1.08	1.79	0.50	0.57	0.50	0.93	3.51	0.79	0.29	0.57	0.43	0.00	1.58	2.44	4.80	40.72	57.56	95.77	2.37	6.88	6.95	2.29	21.58
1203	10.58	3.98	4.95	1.50	1.80	1.73	1.28	1.28	0.98	1.13	1.13	1.13	1.35	0.98	0.75	0.53	0.23	0.68	1.95	4.43	38.26	59.79	93.77	2.78	11.63	6.60	1.05	24.38
1204	47.69	32.45	34.65	26.08	32.98	18.95	20.55	19.41	17.59	13.12	15.24	20.92	12.89	9.55	11.90	13.04	4.85	24.34	27.14	36.69	40.79	47.46	82.64	5.91	4.32	5.00	1.44	28.58
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9471	19.61	4.00	11.19	1.75	2.57	1.75	1.85	0.72	1.13	2.26	1.33	0.92	1.44	0.62	0.82	0.72	0.62	1.03	3.39	5.03	27.82	32.14	93.02	1.03	19.10	11.91	2.87	42.61

Lampiran 2a. Hasil Pengelompokan Kabupaten/Kota dengan Metode K-Means dengan k=2

Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota
1	1101 Simeulue	2	1102 Aceh Singkil
	1103 Aceh Selatan		1105 Aceh Timur
	1104 Aceh Tenggara		1106 Aceh Tengah
	1107 Aceh Barat		1109 Pidie
	1108 Aceh Besar		1110 Bireun
	1112 Aceh Barat Daya		1111 Aceh Utara
	1113 Gayo Lues		1117 Bener Meriah
	1114 Aceh Tamiang		1172 Sabang
	1115 Nagan Raya		1175 Subulussalam
	1116 Aceh Jaya		1201 Nias
	1118 Pidie Jaya		1204 TAPANULI TENGAH
	1171 Banda Aceh		1225 Nias Barat
	1173 Langsa		1278 Gunung Sitoli
	1174 Lhokseumawe		1301 Kep. Mentawai
	1202 Mandailing Natal		1303 Solok
	1203 Tapanuli Selatan		1304 Sijunjung
	1205 Tapanuli Utara		1305 Tanah Datar
	1206 Toba Samosir		1308 Limapuluh kota
	1207 Labuhan Batu		1312 Pasaman Barat
	1208 Asahan		1374 Padang Panjang
	1209 Simalungun		1401 Kuantan Singingi
	1210 Dairi		1572 Sungai Penuh
	1211 Karo		1608 OKU Selatan
	1212 Deli Serdang		1611 Empat Lawang
	1213 Langkat		1671 Palembang
	1214 Nias Selatan		1674 Lubuk Linggau
	1215 Humbang Hasundutan		1701 Bengkulu Selatan
	1216 Pak Pak Bharat		1705 Seluma
	1217 Samosir		1708 Kepahiang
	1218 Serdang Bedagai		1871 Bandar Lampung
	1219 Batu Bara		1902 Belitung
	1220 Padang Lawas Utara		1903 Bangka Barat
	1221 Padang Lawas		1905 Bangka Selatan
	1222 Labuan Batu Selatan		1971 Pangkal Pinang
	1223 Labuan Batu Utara		2105 Kep. Anambas
	1224 Nias Utara		3172 Jakarta Timur
	1271 Sibolga		3175 Jakarta Utara
	1272 Tanjung Balai		3201 Bogor
	1273 Pematang Siantar		3202 Sukabumi

1274	Tebing Tinggi	3203	Cianjur
1275	Medan	3204	Bandung
1276	Binjai	3205	Garut
1277	Padang Sidempuan	3206	Tasikmalaya
1302	Pesisir Selatan	3207	Ciamis
1306	Padang Pariaman	3210	Majalengka
1307	Agam	3211	Sumedang
1309	Pasaman	3213	Subang
1310	Solok Selatan	3214	Purwakarta
1311	Dharmasraya	3217	Bandung Barat
1371	Padang	3271	Bogor
1372	Solok	3272	Sukabumi
1373	Sawah Lunto	3275	Bekasi
1375	Bukittinggi	3276	Depok
1376	Payakumbuh	3277	Cimahi
1377	Pariaman	3278	Tasikmalaya
1402	Indragiri Hulu	3302	Banyumas
1403	Indragiri Hilir	3303	Purbalingga
1404	Pelalawan	3304	Banjarnegara
1405	Siak	3308	Magelang
1406	Kampar	3323	Temanggung
1407	Rokan Hulu	3325	Batang
1408	Bengkalis	3327	Pemalang
1409	Rokan Hilir	3328	Tegal
1410	Kep. Meranti	3373	Salatiga
1471	Pekanbaru	3401	Kulon Progo
1473	Dumai	3403	Gunung Kidul
1501	Kerinci	3471	Yogyakarta
1502	Merangin	3501	Pacitan
1503	Sarolangun	3507	Malang
1504	Batang Hari	3508	Lumajang
1505	Muaro Jambi	3509	Jember
1506	Tanjung Jabung Timur	3510	Banyuwangi
1507	Tanjung Jabung Barat	3512	Situbondo
1508	Tebo	3513	Probolinggo
1509	Bungo	3515	Sidoarjo
1571	Jambi	3525	Gresik
1601	Ogan Komering Ulu	3527	Sampang
1602	Ogan Komering Ilir	3528	Pamekasan
1603	Muara Enim	3529	Sumenep
1604	Lahat	3574	Probolinggo
1605	Musi Rawas	3578	Surabaya
1606	Musi Banyuasin	3601	Pandeglang

1607	Banyu Asin	3603	Tangerang
1609	OKU Timur	3671	Tangerang
1610	Ogan Ilir	3672	Cilegon
1672	Prabumulih	5105	Klungkung
1673	Pagar Alam	5106	Bangli
1702	Rejang Lebong	5202	Lombok Tengah
1703	Bengkulu Utara	5203	Lombok Timur
1704	Kaur	5206	Bima
1706	Mukomuko	5207	Sumbawa Barat
1707	Lebong	5301	Sumba Barat
1709	Bengkulu Tengah	5302	Sumba Timur
1771	Bengkulu	5303	Kupang
1801	Lampung Barat	5305	TTU
1802	Tanggamus	5307	Alor
1803	Lampung Selatan	5308	Lembata
1804	Lampung Timur	5309	Flores Timur
1805	Lampung Tengah	5310	Sikka
1806	Lampung Utara	5311	Ende
1807	Way Kanan	5312	Ngada
1808	Tulangbawang	5314	Rote Ndao
1809	Pesawaran	5315	Manggarai Barat
1810	Pringsewu	5316	Sumba Tengah
1811	Mesuji	5317	Sumba Barat Daya
1812	Tulangbawang Barat	5318	Nagekeo
1872	Metro	5319	Manggarai Timur
1901	Bangka	5320	Sabu Raijua
1904	Bangka Tengah	6104	Mempawah
1906	Belitung Timur	6109	Sekadau
2101	Karimun	6201	Kotawaringin Barat
2102	Bintan	6203	Kapuas
2103	Natuna	6209	Katingan
2104	Lingga	6211	Gunung Mas
2171	Batam	6301	Tanah Laut
2172	Tanjung Pinang	6302	Kota Baru
3101	Kep. Seribu	6303	Banjar
3171	Jakarta Selatan	6304	Barito Kuala
3173	Jakarta Pusat	6310	Tanah Bumbu
3174	Jakarta Barat	6402	Kutai Barat
3208	Kuningan	7101	Bolaang Mongondow
3209	Cirebon	7104	Kep. Talaud
3212	Indramayu	7107	Bolaang Mongondow Utara
3215	Karawang	7108	Siau Tagu Landang Biaro
3216	Bekasi	7110	Bolaang Mongondow Selatan

3273	Bandung	7111	Bolaang Mongondow Timur
3274	Cirebon	7172	Bitung
3279	Banjar	7173	Tomohon
3301	Cilacap	7201	Banggai Kepulauan
3305	Kebumen	7203	Morowali
3306	Purworejo	7204	Poso
3307	Wonosobo	7206	Toli-toli
3309	Boyolali	7208	Parigi Moutong
3310	Klaten	7209	Tojo Una-una
3311	Sukoharjo	7210	Sigi
3312	Wonogiri	7302	Bulukumba
3313	Karanganyar	7303	Bantaeng
3314	Sragen	7304	Jeneponto
3315	Grobogan	7306	Gowa
3316	Blora	7307	Sinjai
3317	Rembang	7310	Barru
3318	Pati	7311	Bone
3319	Kudus	7313	Wajo
3320	Jepara	7315	Pinrang
3321	Demak	7316	Enrekang
3322	Semarang	7317	Luwu
3324	Kendal	7318	Tana Toraja
3326	Pekalongan	7322	Luwu Utara
3329	Brebes	7326	Toraja Utara
3371	Magelang	7371	Makassar
3372	Surakarta	7373	Palopo
3374	Semarang	7404	Kolaka
3375	Pekalongan	7405	Konawe Selatan
3376	Tegal	7472	BauBau
3402	Bantul	7503	Pohuwato
3404	Sleman	7571	Gorontalo
3502	Ponorogo	7602	Polewali Mandar
3503	Trenggalek	7605	Mamuju Utara
3504	Tulungagung	8102	Maluku Tenggara
3505	Blitar	8107	Seram Bagian Timur
3506	Kediri	8108	Maluku Barat daya
3511	Bondowoso	8109	Buru Selatan
3514	Pasuruan	8172	Tual
3516	Mojokerto	8202	Halmahera Tengah
3517	Jombang	8203	Kep. Sula
3518	Nganjuk	8206	Halmahera Timur
3519	Madiun	8207	Pulau Morotai
3520	Magetan	8271	Ternate

3521	Ngawi	9101	Fakfak
3522	Bojonegoro	9408	Kep. Yapen
3523	Tuban	9412	Mimika
3524	Lamongan	9417	Peg. Bintang
3526	Bangkalan	9418	Tolikara
3571	Kediri	9426	Waropen
3572	Blitar	9428	Mamberamo Raya
3573	Malang	9430	Lanny Jaya
3575	Pasuruan	9432	Yalimo
3576	Mojokerto	9435	Intan Jaya
3577	Madiun		
3579	Batu		
3602	Lebak		
3604	Serang		
3673	Serang		
3674	Tangerang Selatan		
5101	Jembrana		
5102	Tabanan		
5103	Badung		
5104	Gianyar		
5107	Karang asem		
5108	Buleleng		
5171	Denpasar		
5201	Lombok Barat		
5204	Sumbawa		
5205	Dompu		
5208	Lombok Utara		
5271	Mataram		
5272	Bima		
5304	TTS		
5306	Belu		
5313	Manggarai		
5371	Kupang		
6101	Sambas		
6102	Bengkayang		
6103	Landak		
6105	Sanggau		
6106	Ketapang		
6107	Sintang		
6108	Kapuas Hulu		
6110	Melawi		
6111	Kayong Utara		
6112	Kubu Raya		

	6171	Pontianak	
	6172	Singkawang	
	6202	Kotawaringin Timur	
	6204	Barito Selatan	
	6205	Barito Utara	
	6206	Sukamara	
	6207	Lamandau	
	6208	Seruyan	
	6210	Pulang Pisau	
	6212	Barito Timur	
	6213	Murung Raya	
	6271	Palangkaraya	
	6305	Tapin	
	6306	Hulu Sungai Selatan	
	6307	Hulu Sungai Tengah	
	6308	Hulu Sungai Utara	
	6309	Tabalong	
	6311	Balangan	
	6371	Banjarmasin	
	6372	Banjar Baru	
	6401	Paser	
	6403	Kutai Kartanegara	
	6404	Kutai Timur	
	6405	Berau	
	6406	Malinau	
	6407	Bulungan	
	6408	Tana Tidung	
	6409	Penajem Paser Utara	
	6410	Nunukan	
	6471	Balikpapan	
	6472	Samarinda	
	6473	Tarakan	
	6474	Bontang	
	7102	Minahasa	
	7103	Kep. Sangihe	
	7105	Minahasa Selatan	
	7106	Minahasa Utara	
	7109	Minahasa Tenggara	
	7171	Manado	
	7174	Kotamobagu	
	7202	Banggai	
	7205	Donggala	
	7207	Buol	

7271	Palu	
7301	Kep. Selayar	
7305	Takalar	
7308	Maros	
7309	Pangkajene dan Kep.	
7312	Soppeng	
7314	Sidenreng Rappang	
7325	Luwu Timur	
7372	Parepare	
7401	Buton	
7402	Muna	
7403	Konawe	
7406	Bombana	
7407	Wakatobi	
7408	Kolaka Utara	
7409	Buton Utara	
7410	Konawe Utara	
7471	Kendari	
7501	Boalemo	
7502	Gorontalo	
7504	Bone Bolango	
7505	Gorontalo Utara	
7601	Majene	
7603	Mamasa	
7604	Mamuju	
8101	Maluku Tenggara Barat	
8103	Maluku Tengah	
8104	Buru	
8105	Kep. Aru	
8106	Seram Bagian Barat	
8171	Ambon	
8201	Halmahera Barat	
8204	Halmahera Selatan	
8205	Halmahera Utara	
8272	Tidore Kepulauan	
9102	Kaimana	
9103	Teluk Wondama	
9104	Teluk Bintuni	
9105	Manokwari	
9106	Sorong Selatan	
9107	Sorong	
9108	Raja Ampat	
9109	Tambrauw	

	9110	Maybrat		
	9171	Sorong		
	9401	Merauke		
	9402	Jayawijaya		
	9403	Jayapura		
	9404	Nabire		
	9409	Biak Numfor		
	9410	Panjai		
	9411	Puncak Jaya		
	9413	Boven Digoel		
	9414	Mappi		
	9415	Asmat		
	9416	Yahukimo		
	9419	Sarmi		
	9420	Keerom		
	9427	Supiori		
	9429	Nduga		
	9431	Mamberamo Tengah		
	9433	Puncak		
	9434	Dogiyai		
	9436	Deiyai		
	9471	Jayapura		

Lampiran 2b. Hasil Pengelompokan Kabupaten/Kota dengan Metode K-Means dengan k=3

Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota
1	1103 Aceh Selatan	2	1102 Aceh Singkil	3	1101 Simeulue
	1107 Aceh Barat		1105 Aceh Timur		1104 Aceh Tenggara
	1108 Aceh Besar		1106 Aceh Tengah		1116 Aceh Jaya
	1109 Pidie		1110 Bireun		1118 Pidie Jaya
	1112 Aceh Barat Daya		1111 Aceh Utara		1202 Mandailing Natal
	1113 Gayo Lues		1117 Bener Meriah		1203 TAPANULI SELATAN
	1114 Aceh Tamiang		1175 Subulussalam		1205 TAPANULI UTARA
	1115 Nagan Raya		1201 Nias		1206 TOBA SAMOSIR
	1171 Banda Aceh		1204 TAPANULI TENGAH		1210 Dairi
	1172 Sabang		1278 Gunung Sitoli		1211 Karo
	1173 Langsa		1304 Sijunjung		1214 Nias Selatan
	1174 Lhokseumawe		1305 Tanah Datar		1215 Humbang Hasundutan
	1207 LABUHAN BATU		1312 Pasaman Barat		1216 Pak Pak Bharat
	1208 ASAHAN		1608 OKU Selatan		1217 Samosir
	1209 Simalungun		1671 Palembang		1220 Padang Lawas Utara
	1212 Deli Serdang		1674 Lubuk Linggau		1224 Nias Utara
	1213 Langkat		1701 Bengkulu Selatan		1225 Nias Barat
	1218 Serdang Bedagai		1871 Bandar Lampung		1301 Kep. Mentawai
	1219 Batu Bara		1902 Belitung		1309 Pasaman
	1221 Padang Lawas		1903 Bangka Barat		1310 Solok Selatan
	1222 Labuan Batu Selatan		1971 Pangkal Pinang		1501 Kerinci
	1223 Labuan Batu Utara		2105 Kep. Anambas		1502 Merangin
	1271 Sibolga		3201 Bogor		1503 Sarolangun
	1272 Tanjung Balai		3202 Sukabumi		1505 Muaro Jambi
	1273 Pematang Siantar		3203 Cianjur		1506 Tanjung Jabung Timur
	1274 Tebing Tinggi		3204 Bandung		1508 Tebo
	1275 Medan		3206 Tasikmalaya		1602 Ogan Komering Ilir
	1276 Binjai		3207 Ciamis		1603 Muara Enim
	1277 Padang Sidempuan		3211 Sumedang		1604 Lahat
	1302 Pesisir Selatan		3213 Subang		1605 Musi Rawas
	1303 Solok		3214 Purwakarta		1606 Musi Banyuasin
	1306 Padang Pariaman		3217 Bandung Barat		1607 Banyu Asin
	1307 Agam		3271 Bogor		1609 OKU Timur
	1308 Limapuluh kota		3272 Sukabumi		1610 Ogan Ilir
	1311 Dharmasraya		3275 Bekasi		1611 Empat Lawang
	1371 Padang		3276 Depok		1673 Pagar Alam
	1372 Solok		3277 Cimahi		1703 Bengkulu Utara
	1373 Sawah Lunto		3278 Tasikmalaya		1704 Kaur
	1374 Padang Panjang				

1375	Bukittinggi	3303	Purbalingga	1705	Seluma
1376	Payakumbuh	3308	Magelang	1707	Lebong
1377	Pariaman	3323	Temanggung	1708	Kepahiang
1401	Kuantan Singingi	3325	Batang	1709	Bengkulu Tengah
1402	Indragiri Hulu	3327	Pemalang	1801	Lampung Barat
1403	Indragiri Hilir	3328	Tegal	1804	Lampung Timur
1404	Pelalawan	3401	Kulon Progo	1805	Lampung Tengah
1405	Siak	3403	Gunung Kidul	1807	Way Kanan
1406	Kampar	3471	Yogyakarta	1808	Tulangbawang
1407	Rokan Hulu	3507	Malang	1809	Pesawaran
1408	Bengkalis	3508	Lumajang	1812	Tulangbawang Barat
1409	Rokan Hilir	3509	Jember	1905	Bangka Selatan
1410	Kep. Meranti	3510	Banyuwangi	3209	Cirebon
1471	Pekanbaru	3512	Situbondo	3305	Kebumen
1473	Dumai	3513	Probolinggo	3313	Karanganyar
1504	Batang Hari	3527	Sampang	3316	Blora
1507	Tanjung Jabung Barat	3528	Pamekasan	3324	Kendal
1509	Bungo	3529	Sumenep	3501	Pacitan
1571	Jambi	3574	Probolinggo	3503	Trenggalek
1572	Sungai Penuh	3578	Surabaya	3511	Bondowoso
1601	Ogan Komering Ulu	3601	Pandeglang	3518	Nganjuk
1672	Prabumulih	3672	Cilegon	3522	Bojonegoro
1702	Rejang Lebong	5106	Bangli	3523	Tuban
1706	Mukomuko	5202	Lombok Tengah	3524	Lamongan
1771	Bengkulu	5203	Lombok Timur	5104	Gianyar
1802	Tanggamus	5206	Bima	5107	Karang asem
1803	Lampung Selatan	5207	Sumbawa Barat	5201	Lombok Barat
1806	Lampung Utara	5301	Sumba Barat	5208	Lombok Utara
1810	Pringsewu	5302	Sumba Timur	5303	Kupang
1811	Mesuji	5305	TTU	5304	TTS
1872	Metro	5307	Alor	5306	Belu
1901	Bangka	5308	Lembata	5309	Flores Timur
1904	Bangka Tengah	5310	Sikka	5313	Manggarai
1906	Belitung Timur	5311	Ende	5320	Sabu Raijua
2101	Karimun	5312	Ngada	6101	Sambas
2102	Bintan	5314	Rote Ndao	6102	Bengkayang
2103	Natuna	5315	Manggarai Barat	6103	Landak
2104	Lingga	5316	Sumba Tengah	6106	Ketapang
2171	Batam	5317	Sumba Barat Daya	6107	Sintang
2172	Tanjung Pinang	5318	Nagekeo	6108	Kapuas Hulu
3101	Kep. Seribu	5319	Manggarai Timur	6110	Melawi
3171	Jakarta Selatan	6104	Mempawah	6213	Murung Raya
3172	Jakarta Timur	6109	Sekadau	6372	Banjar Baru

3173	Jakarta Pusat	6209	Katingan	6404	Kutai Timur
3174	Jakarta Barat	6211	Gunung Mas	7205	Donggala
3175	Jakarta Utara	6301	Tanah Laut	7207	Buol
3205	Garut	6303	Banjar	7316	Enrekang
3208	Kuningan	6310	Tanah Bumbu	7325	Luwu Timur
3210	Majalengka	6402	Kutai Barat	7401	Buton
3212	Indramayu	7101	Bolaang Mongondow	7402	Muna
3215	Karawang	7104	Kep. Talaud	7408	Kolaka Utara
3216	Bekasi	7107	Bolaang M. Utara	7505	Gorontalo Utara
3273	Bandung	7108	Siau Tagu L. Biaro	7601	Majene
3274	Cirebon	7110	Bolaang M. Selatan	7603	Mamasa
3279	Banjar	7111	Bolaang M. Timur	7604	Mamuju
3301	Cilacap	7172	Bitung	7605	Mamuju Utara
3302	Banyumas	7201	Banggai Kepulauan	8101	Maluku Tenggara Barat
3304	Banjarnegara	7203	Morowali	8105	Kep. Aru
3306	Purworejo	7204	Poso	8106	Seram Bagian Barat
3307	Wonosobo	7206	Toli-toli	8109	Buru Selatan
3309	Boyolali	7208	Parigi Moutong	8204	Halmahera Selatan
3310	Klaten	7209	Tojo Una-una	8205	Halmahera Utara
3311	Sukoharjo	7210	Sigi	9102	Kaimana
3312	Wonogiri	7302	Bulukumba	9103	Teluk Wondama
3314	Sragen	7303	Bantaeng	9104	Teluk Bintuni
3315	Grobogan	7304	Jeneponto	9106	Sorong Selatan
3317	Rembang	7306	Gowa	9107	Sorong
3318	Pati	7307	Sinjai	9109	Tambrauw
3319	Kudus	7310	Barru	9110	Maybrat
3320	Jepara	7311	Bone	9402	Jayawijaya
3321	Demak	7313	Wajo	9403	Jayapura
3322	Semarang	7315	Pinrang	9409	Biak Numfor
3326	Pekalongan	7317	Luwu	9410	Panjai
3329	Brebes	7318	Tana Toraja	9411	Puncak Jaya
3371	Magelang	7322	Luwu Utara	9413	Boven Digoel
3372	Surakarta	7326	Toraja Utara	9414	Mappi
3373	Salatiga	7371	Makassar	9415	Asmat
3374	Semarang	7373	Palopo	9416	Yahukimo
3375	Pekalongan	7404	Kolaka	9417	Peg. Bintang
3376	Tegal	7405	Konawe Selatan	9427	Supiori
3402	Bantul	7472	BauBau	9429	Nduga
3404	Sleman	7503	Pohuwato	9431	Mamberamo Tengah
3502	Ponorogo	7571	Gorontalo	9432	Yalimo
3504	Tulungagung	7602	Polewali Mandar	9433	Puncak
3505	Blitar	8102	Maluku Tenggara	9434	Dogiyai
3506	Kediri	8107	Seram Bagian Timur	9435	Intan Jaya

3514	Pasuruan	8108	Maluku Barat daya	9436	Deiyai
3515	Sidoarjo	8172	Tual		
3516	Mojokerto	8202	Halmahera Tengah		
3517	Jombang	8203	Kep. Sula		
3519	Madiun	8206	Halmahera Timur		
3520	Magetan	8207	Pulau Morotai		
3521	Ngawi	8271	Ternate		
3525	Gresik	9408	Kep. Yapen		
3526	Bangkalan	9412	Mimika		
3571	Kediri	9418	Tolikara		
3572	Blitar	9428	Mamberamo Raya		
3573	Malang	9430	Lanny Jaya		
3575	Pasuruan				
3576	Mojokerto				
3577	Madiun				
3579	Batu				
3602	Lebak				
3603	Tangerang				
3604	Serang				
3671	Tangerang				
3673	Serang				
3674	Tangerang Selatan				
5101	Jembrana				
5102	Tabanan				
5103	Badung				
5105	Klungkung				
5108	Buleleng				
5171	Denpasar				
5204	Sumbawa				
5205	Dompu				
5271	Mataram				
5272	Bima				
5371	Kupang				
6105	Sanggau				
6111	Kayong Utara				
6112	Kubu Raya				
6171	Pontianak				
6172	Singkawang				
6201	Kotawaringin Barat				
6202	Kotawaringin Timur				
6203	Kapuas				
6204	Barito Selatan				
6205	Barito Utara				

6206	Sukamara		
6207	Lamandau		
6208	Seruyan		
6210	Pulang Pisau		
6212	Barito Timur		
6271	Palangkaraya		
6302	Kota Baru		
6304	Barito Kuala		
6305	Tapin		
6306	Hulu Sungai Selatan		
6307	Hulu Sungai Tengah		
6308	Hulu Sungai Utara		
6309	Tabalong		
6311	Balangan		
6371	Banjarmasin		
6401	Paser		
6403	Kutai Kartanegara		
6405	Berau		
6406	Malinau		
6407	Bulungan		
6408	Tana Tidung		
6409	Penajem Paser Utara		
6410	Nunukan		
6471	Balikpapan		
6472	Samarinda		
6473	Tarakan		
6474	Bontang		
7102	Minahasa		
7103	Kep. Sangihe		
7105	Minahasa Selatan		
7106	Minahasa Utara		
7109	Minahasa Tenggara		
7171	Manado		
7173	Tomohon		
7174	Kotamobagu		
7202	Banggai		
7271	Palu		
7301	Kep. Selayar		
7305	Takalar		
7308	Maros		
7309	Pangkajene dan Kep.		
7312	Soppeng		
7314	Sidenreng Rappang		

7372	Parepare				
7403	Konawe				
7406	Bombana				
7407	Wakatobi				
7409	Buton Utara				
7410	Konawe Utara				
7471	Kendari				
7501	Boalemo				
7502	Gorontalo				
7504	Bone Bolango				
8103	Maluku Tengah				
8104	Buru				
8171	Ambon				
8201	Halmahera Barat				
8272	Tidore Kepulauan				
9101	Fakfak				
9105	Manokwari				
9108	Raja Ampat				
9171	Sorong				
9401	Merauke				
9404	Nabire				
9419	Sarmi				
9420	Keerom				
9426	Waropen				
9471	Jayapura				

Lampiran 2c. Hasil Pengelompokan Kabupaten/Kota dengan Metode Kernel

K-Means dengan k=2

Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota
1	1103 Aceh Selatan	2	1101 Simeulue
	1104 Aceh Tenggara		1102 Aceh Singkil
	1108 Aceh Besar		1105 Aceh Timur
	1112 Aceh Barat Daya		1106 Aceh Tengah
	1113 Gayo Lues		1107 Aceh Barat
	1114 Aceh Tamiang		1109 Pidie
	1171 Banda Aceh		1110 Bireun
	1173 Langsa		1111 Aceh Utara
	1174 Lhokseumawe		1115 Nagan Raya
	1202 Mandailing Natal		1116 Aceh Jaya
	1203 TAPANULI SELATAN		1117 Bener Meriah
	1205 TAPANULI UTARA		1118 Pidie Jaya
	1206 TOBA SAMOSIR		1172 Sabang
	1207 LABUHAN BATU		1175 Subulussalam
	1208 ASAHAN		1201 Nias
	1209 Simalungun		1204 TAPANULI TENGAH
	1210 Dairi		1214 Nias Selatan
	1211 Karo		1216 Pak Pak Bharat
	1212 Deli Serdang		1225 Nias Barat
	1213 Langkat		1278 Gunung Sitoli
	1215 Humbang Hasundutan		1301 Kep. Mentawai
	1217 Samosir		1303 Solok
	1218 Serdang Bedagai		1304 Sijunjung
	1219 Batu Bara		1305 Tanah Datar
	1220 Padang Lawas Utara		1308 Limapuluh kota
	1221 Padang Lawas		1309 Pasaman
	1222 Labuan Batu Selatan		1312 Pasaman Barat
	1223 Labuan Batu Utara		1374 Padang Panjang
	1224 Nias Utara		1401 Kuantan Singingi
	1271 Sibolga		1403 Indragiri Hilir
	1272 Tanjung Balai		1506 Tanjung Jabung Timur
	1273 Pematang Siantar		1572 Sungai Penuh
	1274 Tebing Tinggi		1608 OKU Selatan
	1275 Medan		1611 Empat Lawang
	1276 Binjai		1671 Palembang
	1277 Padang Sidempuan		1673 Pagar Alam
	1302 Pesisir Selatan		1674 Lubuk Linggau
	1306 Padang Pariaman		1701 Bengkulu Selatan
	1307 Agam		1704 Kaur

1310	Solok Selatan	1705	Seluma
1311	Dharmasraya	1707	Lebong
1371	Padang	1708	Kepahiang
1372	Solok	1709	Bengkulu Tengah
1373	Sawah Lunto	1801	Lampung Barat
1375	Bukittinggi	1806	Lampung Utara
1376	Payakumbuh	1810	Pringsewu
1377	Pariaman	1871	Bandar Lampung
1402	Indragiri Hulu	1901	Bangka
1404	Pelalawan	1902	Belitung
1405	Siak	1903	Bangka Barat
1406	Kampar	1905	Bangka Selatan
1407	Rokan Hulu	1971	Pangkal Pinang
1408	Bengkalis	2103	Natuna
1409	Rokan Hilir	2104	Lingga
1410	Kep. Meranti	2105	Kep. Anambas
1471	Pekanbaru	3172	Jakarta Timur
1473	Dumai	3175	Jakarta Utara
1501	Kerinci	3201	Bogor
1502	Merangin	3202	Sukabumi
1503	Sarolangun	3203	Cianjur
1504	Batang Hari	3204	Bandung
1505	Muaro Jambi	3205	Garut
1507	Tanjung Jabung Barat	3206	Tasikmalaya
1508	Tebo	3207	Ciamis
1509	Bungo	3210	Majalengka
1571	Jambi	3211	Sumedang
1601	Ogan Komering Ulu	3213	Subang
1602	Ogan Komering Ilir	3214	Purwakarta
1603	Muara Enim	3217	Bandung Barat
1604	Lahat	3271	Bogor
1605	Musi Rawas	3272	Sukabumi
1606	Musi Banyuasin	3275	Bekasi
1607	Banyu Asin	3276	Depok
1609	OKU Timur	3277	Cimahi
1610	Ogan Ilir	3278	Tasikmalaya
1672	Prabumulih	3302	Banyumas
1702	Rejang Lebong	3303	Purbalingga
1703	Bengkulu Utara	3304	Banjarnegara
1706	Mukomuko	3305	Kebumen
1771	Bengkulu	3307	Wonosobo
1802	Tanggamus	3308	Magelang
1803	Lampung Selatan	3309	Boyolali

1804	Lampung Timur	3312	Wonogiri
1805	Lampung Tengah	3322	Semarang
1807	Way Kanan	3323	Temanggung
1808	Tulangbawang	3325	Batang
1809	Pesawaran	3326	Pekalongan
1811	Mesuji	3327	Pemalang
1812	Tulangbawang Barat	3328	Tegal
1872	Metro	3329	Brebes
1904	Bangka Tengah	3372	Surakarta
1906	Belitung Timur	3373	Salatiga
2101	Karimun	3401	Kulon Progo
2102	Bintan	3403	Gunung Kidul
2171	Batam	3471	Yogyakarta
2172	Tanjung Pinang	3501	Pacitan
3101	Kep. Seribu	3502	Ponorogo
3171	Jakarta Selatan	3507	Malang
3173	Jakarta Pusat	3508	Lumajang
3174	Jakarta Barat	3509	Jember
3208	Kuningan	3510	Banyuwangi
3209	Cirebon	3511	Bondowoso
3212	Indramayu	3512	Situbondo
3215	Karawang	3513	Probolinggo
3216	Bekasi	3515	Sidoarjo
3273	Bandung	3520	Magetan
3274	Cirebon	3521	Ngawi
3279	Banjar	3522	Bojonegoro
3301	Cilacap	3523	Tuban
3306	Purworejo	3525	Gresik
3310	Klaten	3527	Sampang
3311	Sukoharjo	3528	Pamekasan
3313	Karanganyar	3529	Sumenep
3314	Sragen	3574	Probolinggo
3315	Grobogan	3578	Surabaya
3316	Blora	3579	Batu
3317	Rembang	3601	Pandeglang
3318	Pati	3603	Tangerang
3319	Kudus	3671	Tangerang
3320	Jepara	3672	Cilegon
3321	Demak	5105	Klungkung
3324	Kendal	5106	Bangli
3371	Magelang	5107	Karang asem
3374	Semarang	5201	Lombok Barat
3375	Pekalongan	5202	Lombok Tengah

3376	Tegal	5203	Lombok Timur
3402	Bantul	5206	Bima
3404	Sleman	5207	Sumbawa Barat
3503	Trenggalek	5208	Lombok Utara
3504	Tulungagung	5301	Sumba Barat
3505	Blitar	5302	Sumba Timur
3506	Kediri	5303	Kupang
3514	Pasuruan	5304	TTS
3516	Mojokerto	5305	TTU
3517	Jombang	5307	Alor
3518	Nganjuk	5308	Lembata
3519	Madiun	5309	Flores Timur
3524	Lamongan	5310	Sikka
3526	Bangkalan	5311	Ende
3571	Kediri	5312	Ngada
3572	Blitar	5314	Rote Ndao
3573	Malang	5315	Manggarai Barat
3575	Pasuruan	5316	Sumba Tengah
3576	Mojokerto	5317	Sumba Barat Daya
3577	Madiun	5318	Nagekeo
3602	Lebak	5319	Manggarai Timur
3604	Serang	5320	Sabu Raijua
3673	Serang	6101	Sambas
3674	Tangerang Selatan	6102	Bengkayang
5101	Jembrana	6103	Landak
5102	Tabanan	6104	Mempawah
5103	Badung	6106	Ketapang
5104	Gianyar	6109	Sekadau
5108	Buleleng	6110	Melawi
5171	Denpasar	6111	Kayong Utara
5204	Sumbawa	6201	Kotawaringin Barat
5205	Dompu	6203	Kapuas
5271	Mataram	6209	Katingan
5272	Bima	6211	Gunung Mas
5306	Belu	6301	Tanah Laut
5313	Manggarai	6302	Kota Baru
5371	Kupang	6303	Banjar
6105	Sanggau	6304	Barito Kuala
6107	Sintang	6310	Tanah Bumbu
6108	Kapuas Hulu	6402	Kutai Barat
6112	Kubu Raya	7101	Bolaang Mongondow
6171	Pontianak	7104	Kep. Talaud
6172	Singkawang	7107	Bolaang Mongondow Utara

6202	Kotawaringin Timur	7108	Siau Tagu Landang Biaro
6204	Barito Selatan	7109	Minahasa Tenggara
6205	Barito Utara	7110	Bolaang Mongondow Selatan
6206	Sukamara	7111	Bolaang Mongondow Timur
6207	Lamandau	7172	Bitung
6208	Seruyan	7173	Tomohon
6210	Pulang Pisau	7201	Banggai Kepulauan
6212	Barito Timur	7203	Morowali
6213	Murung Raya	7204	Poso
6271	Palangkaraya	7206	Toli-toli
6305	Tapin	7207	Buol
6306	Hulu Sungai Selatan	7208	Parigi Moutong
6307	Hulu Sungai Tengah	7209	Tojo Una-una
6308	Hulu Sungai Utara	7210	Sigi
6309	Tabalong	7302	Bulukumba
6311	Balangan	7303	Bantaeng
6371	Banjarmasin	7304	Jeneponto
6372	Banjar Baru	7306	Gowa
6401	Paser	7307	Sinjai
6403	Kutai Kartanegara	7310	Barru
6404	Kutai Timur	7311	Bone
6405	Berau	7313	Wajo
6406	Malinau	7315	Pinrang
6407	Bulungan	7316	Enrekang
6408	Tana Tidung	7317	Luwu
6409	Penajem Paser Utara	7318	Tana Toraja
6410	Nunukan	7322	Luwu Utara
6471	Balikpapan	7325	Luwu Timur
6472	Samarinda	7326	Toraja Utara
6473	Tarakan	7371	Makassar
6474	Bontang	7373	Palopo
7102	Minahasa	7401	Buton
7103	Kep. Sangihe	7403	Konawe
7105	Minahasa Selatan	7404	Kolaka
7106	Minahasa Utara	7405	Konawe Selatan
7171	Manado	7472	BauBau
7174	Kotamobagu	7503	Pohuwato
7202	Banggai	7505	Gorontalo Utara
7205	Donggala	7571	Gorontalo
7271	Palu	7602	Polewali Mandar
7301	Kep. Selayar	7603	Mamasa
7305	Takalar	7605	Mamuju Utara
7308	Maros	8102	Maluku Tenggara

7309	Pangkajene dan Kep.	8105	Kep. Aru
7312	Soppeng	8106	Seram Bagian Barat
7314	Sidenreng Rappang	8107	Seram Bagian Timur
7372	Parepare	8108	Maluku Barat daya
7402	Muna	8109	Buru Selatan
7406	Bombana	8172	Tual
7407	Wakatobi	8202	Halmahera Tengah
7408	Kolaka Utara	8203	Kep. Sula
7409	Buton Utara	8204	Halmahera Selatan
7410	Konawe Utara	8206	Halmahera Timur
7471	Kendari	8207	Pulau Morotai
7501	Boalemo	8271	Ternate
7502	Gorontalo	9101	Fakfak
7504	Bone Bolango	9102	Kaimana
7601	Majene	9106	Sorong Selatan
7604	Mamuju	9408	Kep. Yapen
8101	Maluku Tenggara Barat	9411	Puncak Jaya
8103	Maluku Tengah	9412	Mimika
8104	Buru	9414	Mappi
8171	Ambon	9416	Yahukimo
8201	Halmahera Barat	9417	Peg. Bintang
8205	Halmahera Utara	9418	Tolikara
8272	Tidore Kepulauan	9426	Waropen
9103	Teluk Wondama	9427	Supiori
9104	Teluk Bintuni	9428	Mamberamo Raya
9105	Manokwari	9429	Nduga
9107	Sorong	9430	Lanny Jaya
9108	Raja Ampat	9431	Mamberamo Tengah
9109	Tambrauw	9432	Yalimo
9110	Maybrat	9433	Puncak
9171	Sorong	9435	Intan Jaya
9401	Merauke	9436	Deiyai
9402	Jayawijaya		
9403	Jayapura		
9404	Nabire		
9409	Biak Numfor		
9410	Panjai		
9413	Boven Digoel		
9415	Asmat		
9419	Sarmi		
9420	Keerom		
9434	Dogiyai		
9471	Jayapura		

Lampiran 2d. Hasil Pengelompokan Kabupaten/Kota dengan Metode Kernel

K-Means dengan k=3

Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota
1	1102 Aceh Singkil	2	1101 Simeulue	3	1103 Aceh Selatan
	1105 Aceh Timur		1106 Aceh Tengah		1104 Aceh Tenggara
	1107 Aceh Barat		1110 Bireun		1108 Aceh Besar
	1117 Bener Meriah		1111 Aceh Utara		1109 Pidie
	1175 Subulussalam		1116 Aceh Jaya		1112 Aceh Barat Daya
	1201 Nias		1118 Pidie Jaya		1113 Gayo Lues
	1204 TAPANULI TENGAH		1209 Simalungun		1114 Aceh Tamiang
	1225 Nias Barat		1212 Deli Serdang		1115 Nagan Raya
	1278 Gunung Sitoli		1214 Nias Selatan		1171 Banda Aceh
	1374 Padang Panjang		1216 Pak Pak Bharat		1172 Sabang
	1403 Indragiri Hilir		1301 Kep. Mentawai		1173 Langsa
	1507 Tanjung Jabung Barat		1303 Solok		1174 Lhokseumawe
	1608 OKU Selatan		1304 Sijunjung		1202 Mandailing Natal
	1671 Palembang		1305 Tanah Datar		1203 TAPANULI SELATAN
	1673 Pagar Alam		1307 Agam		1205 TAPANULI UTARA
	1801 Lampung Barat		1308 Limapuluh kota		1206 TOBA SAMOSIR
	1806 Lampung Utara		1309 Pasaman		1207 LABUHAN BATU
	1810 Pringsewu		1310 Solok Selatan		1208 ASAHAN
	1871 Bandar Lampung		1312 Pasaman Barat		1210 Dairi
	1971 Pangkal Pinang		1401 Kuantan Singingi		1211 Karo
	2104 Lingga		1506 Tanjung Jabung Timur		1213 Langkat
	2105 Kep. Anambas		1572 Sungai Penuh		1215 Humbang Hasundutan
	3171 Jakarta Selatan		1603 Muara Enim		1217 Samosir
	3172 Jakarta Timur		1605 Musi Rawas		1218 Serdang Bedagai
	3201 Bogor		1611 Empat Lawang		1219 Batu Bara
	3203 Cianjur		1674 Lubuk Linggau		1220 Padang Lawas Utara
	3204 Bandung		1701 Bengkulu Selatan		1221 Padang Lawas
	3205 Garut		1704 Kaur		1222 Labuan Batu Selatan
	3206 Tasikmalaya		1705 Seluma		1223 Labuan Batu Utara
	3207 Ciamis		1708 Kepahiang		1224 Nias Utara
	3210 Majalengka		1709 Bengkulu Tengah		1271 Sibolga
	3211 Sumedang		1901 Bangka		1272 Tanjung Balai
	3213 Subang		1902 Belitung		1273 Pematang Siantar
	3214 Purwakarta		1903 Bangka Barat		1274 Tebing Tinggi
	3217 Bandung Barat		1905 Bangka Selatan		1275 Medan
	3271 Bogor		1906 Belitung Timur		1276 Binjai
	3272 Sukabumi		3202 Sukabumi		1277 Padang Sidempuan
	3273 Bandung		3403 Gunung Kidul		1302 Pesisir Selatan

3275	Bekasi	3503	Trenggalek	1306	Padang Pariaman
3276	Depok	3510	Banyuwangi	1311	Dharmasraya
3277	Cimahi	3527	Sampang	1371	Padang
3278	Tasikmalaya	3528	Pamekasan	1372	Solok
3302	Banyumas	5105	Klungkung	1373	Sawah Lunto
3303	Purbalingga	5107	Karang asem	1375	Bukittinggi
3304	Banjarnegara	5201	Lombok Barat	1376	Payakumbuh
3305	Kebumen	5203	Lombok Timur	1377	Pariaman
3307	Wonosobo	5206	Bima	1402	Indragiri Hulu
3308	Magelang	5207	Sumbawa Barat	1404	Pelalawan
3309	Boyolali	5208	Lombok Utara	1405	Siak
3310	Klaten	5301	Sumba Barat	1406	Kampar
3312	Wonogiri	5303	Kupang	1407	Rokan Hulu
3313	Karanganyar	5304	TTS	1408	Bengkalis
3322	Semarang	5305	TTU	1409	Rokan Hilir
3323	Temanggung	5307	Alor	1410	Kep. Meranti
3325	Batang	5309	Flores Timur	1471	Pekanbaru
3326	Pekalongan	5310	Sikka	1473	Dumai
3327	Pemalang	5311	Ende	1501	Kerinci
3328	Tegal	5314	Rote Ndao	1502	Merangin
3329	Brebes	5320	Sabu Raijua	1503	Sarolangun
3371	Magelang	6101	Sambas	1504	Batang Hari
3372	Surakarta	6102	Bengkayang	1505	Muaro Jambi
3373	Salatiga	6103	Landak	1508	Tebo
3374	Semarang	6110	Melawi	1509	Bungo
3401	Kulon Progo	6213	Murung Raya	1571	Jambi
3402	Bantul	6302	Kota Baru	1601	Ogan Komering Ulu
3471	Yogyakarta	7101	Bolaang Mongondow	1602	Ogan Komering Ilir
3501	Pacitan	7107	Bolaang Mongondow Utara	1604	Lahat
3502	Ponorogo	7109	Minahasa Tenggara	1606	Musi Banyuasin
3507	Malang	7201	Banggai Kepulauan	1607	Banyu Asin
3508	Lumajang	7203	Morowali	1609	OKU Timur
3509	Jember	7204	Poso	1610	Ogan Ilir
3511	Bondowoso	7205	Donggala	1672	Prabumulih
3512	Situbondo	7207	Buol	1702	Rejang Lebong
3513	Probolinggo	7302	Bulukumba	1703	Bengkulu Utara
3515	Sidoarjo	7303	Bantaeng	1706	Mukomuko
3520	Magetan	7305	Takalar	1707	Lebong
3521	Ngawi	7306	Gowa	1771	Bengkulu
3525	Gresik	7307	Sinjai	1802	Tanggamus
3529	Sumenep	7308	Maros	1803	Lampung Selatan
3574	Probolinggo	7311	Bone	1804	Lampung Timur
3575	Pasuruan			1805	Lampung Tengah

3576	Mojokerto	7313	Wajo	1807	Way Kanan
3577	Madiun	7316	Enrekang	1808	Tulangbawang
3578	Surabaya	7322	Luwu Utara	1809	Pesawaran
3579	Batu	7325	Luwu Timur	1811	Mesuji
3601	Pandeglang	7401	Buton	1812	Tulangbawang Barat
3603	Tangerang	7402	Muna	1872	Metro
3671	Tangerang	7404	Kolaka	1904	Bangka Tengah
3672	Cilegon	7405	Konawe Selatan	2101	Karimun
5106	Bangli	7408	Kolaka Utara	2102	Bintan
5202	Lombok Tengah	7472	BauBau	2103	Natuna
5302	Sumba Timur	7503	Pohuwato	2171	Batam
5308	Lembata	7505	Gorontalo Utara	2172	Tanjung Pinang
5312	Ngada	7601	Majene	3101	Kep. Seribu
5315	Manggarai Barat	7604	Mamuju	3173	Jakarta Pusat
5316	Sumba Tengah	7605	Mamuju Utara	3174	Jakarta Barat
5317	Sumba Barat Daya	8102	Maluku Tenggara	3175	Jakarta Utara
5318	Nagekeo	8105	Kep. Aru	3208	Kuningan
5319	Manggarai Timur	8107	Seram Bagian Timur	3209	Cirebon
6104	Mempawah	8108	Maluku Barat daya	3212	Indramayu
6109	Sekadau	8202	Halmahera Tengah	3215	Karawang
6111	Kayong Utara	8203	Kep. Sula	3216	Bekasi
6203	Kapuas	8204	Halmahera Selatan	3274	Cirebon
6209	Katingan	8206	Halmahera Timur	3279	Banjar
6211	Gunung Mas	9102	Kaimana	3301	Cilacap
6301	Tanah Laut	9106	Sorong Selatan	3306	Purworejo
6303	Banjar	9408	Kep. Yapen	3311	Sukoharjo
6304	Barito Kuala	9412	Mimika	3314	Sragen
6310	Tanah Bumbu	9414	Mappi	3315	Grobogan
6402	Kutai Barat	9417	Peg. Bintang	3316	Blora
7104	Kep. Talaud	9426	Waropen	3317	Rembang
7108	Siau Tagu Landang Biaro	9427	Supiori	3318	Pati
7110	Bolaang Mongondow Selatan	9432	Yalimo	3319	Kudus
7111	Bolaang Mongondow Timur			3320	Jepara
7172	Bitung			3321	Demak
7173	Tomohon			3324	Kendal
7206	Toli-toli			3375	Pekalongan
7208	Parigi Moutong			3376	Tegal
7209	Tojo Una-una			3404	Sleman
7210	Sigi			3504	Tulungagung
7304	Jeneponto			3505	Blitar
7310	Barru			3506	Kediri
				3514	Pasuruan
				3516	Mojokerto

7315	Pinrang			3517	Jombang
7317	Luwu			3518	Nganjuk
7318	Tana Toraja			3519	Madiun
7326	Toraja Utara			3522	Bojonegoro
7371	Makassar			3523	Tuban
7373	Palopo			3524	Lamongan
7504	Bone Bolango			3526	Bangkalan
7571	Gorontalo			3571	Kediri
7602	Polewali Mandar			3572	Blitar
7603	Mamasa			3573	Malang
8109	Buru Selatan			3602	Lebak
8171	Ambon			3604	Serang
8172	Tual			3673	Serang
8207	Pulau Morotai			3674	Tangerang Selatan
8271	Ternate			5101	Jembrana
9101	Fakfak			5102	Tabanan
9171	Sorong			5103	Badung
9411	Puncak Jaya			5104	Gianyar
9416	Yahukimo			5108	Buleleng
9418	Tolikara			5171	Denpasar
9428	Mamberamo Raya			5204	Sumbawa
9429	Nduga			5205	Dompus
9430	Lanny Jaya			5271	Mataram
9431	Mamberamo Tengah			5272	Bima
9433	Puncak			5306	Belu
9435	Intan Jaya			5313	Manggarai
				5371	Kupang
				6105	Sanggau
				6106	Ketapang
				6107	Sintang
				6108	Kapuas Hulu
				6112	Kubu Raya
				6171	Pontianak
				6172	Singkawang
				6201	Kotawaringin Barat
				6202	Kotawaringin Timur
				6204	Barito Selatan
				6205	Barito Utara
				6206	Sukamara
				6207	Lamandau
				6208	Seruyan
				6210	Pulang Pisau
				6212	Barito Timur

				6271	Palangkaraya
				6305	Tapin
				6306	Hulu Sungai Selatan
				6307	Hulu Sungai Tengah
				6308	Hulu Sungai Utara
				6309	Tabalong
				6311	Balangan
				6371	Banjarmasin
				6372	Banjar Baru
				6401	Paser
				6403	Kutai Kartanegara
				6404	Kutai Timur
				6405	Berau
				6406	Malinau
				6407	Bulungan
				6408	Tana Tidung
				6409	Penajem Paser Utara
				6410	Nunukan
				6471	Balikpapan
				6472	Samarinda
				6473	Tarakan
				6474	Bontang
				7102	Minahasa
				7103	Kep. Sangihe
				7105	Minahasa Selatan
				7106	Minahasa Utara
				7171	Manado
				7174	Kotamobagu
				7202	Banggai
				7271	Palu
				7301	Kep. Selayar
				7309	Pangkajene dan Kep.
				7312	Soppeng
				7314	Sidenreng Rappang
				7372	Parepare
				7403	Konawe
				7406	Bombana
				7407	Wakatobi
				7409	Buton Utara
				7410	Konawe Utara
				7471	Kendari
				7501	Boalemo
				7502	Gorontalo

				8101	Maluku Tenggara Barat
				8103	Maluku Tengah
				8104	Buru
				8106	Seram Bagian Barat
				8201	Halmahera Barat
				8205	Halmahera Utara
				8272	Tidore Kepulauan
				9103	Teluk Wondama
				9104	Teluk Bintuni
				9105	Manokwari
				9107	Sorong
				9108	Raja Ampat
				9109	Tambrauw
				9110	Maybrat
				9401	Merauke
				9402	Jayawijaya
				9403	Jayapura
				9404	Nabire
				9409	Biak Numfor
				9410	Panjai
				9413	Boven Digoel
				9415	Asmat
				9419	Sarmi
				9420	Keerom
				9434	Dogiyai
				9436	Deiyai
				9471	Jayapura

**Lampiran 2e. Hasil Pengelompokan Kabupaten/Kota menurut Faktor
Kejiwaan dengan Metode Kernel K-Means dengan k=3**

Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota	Kel	Kode dan Nama Kab/Kota
1	1108 Aceh Besar	2	1102 Aceh Singkil	3	1101 Simeulue
	1112 Aceh Barat Daya		1105 Aceh Timur		1103 Aceh Selatan
	1113 Gayo Lues		1106 Aceh Tengah		1104 Aceh Tenggara
	1114 Aceh Tamiang		1107 Aceh Barat		1173 Langsa
	1115 Nagan Raya		1109 Pidie		1203 TAPANULI SELATAN
	1118 Pidie Jaya		1110 Bireun		1206 TOBA SAMOSIR
	1171 Banda Aceh		1111 Aceh Utara		1207 LABUHAN BATU
	1202 Mandailing Natal		1116 Aceh Jaya		1210 Dairi
	1205 TAPANULI UTARA		1117 Bener Meriah		1215 Humbang Hasundutan
	1211 Karo		1172 Sabang		1217 Samosir
	1213 Langkat		1174 Lhokseumawe		1219 Batu Bara
	1216 Pak Pak Bharat		1175 Subulussalam		1220 Padang Lawas Utara
	1218 Serdang Bedagai		1201 Nias		1222 Labuan Batu Selatan
	1221 Padang Lawas		1204 TAPANULI TENGAH		1223 Labuan Batu Utara
	1273 Pematang Siantar		1208 ASAHAN		1224 Nias Utara
	1274 Tebing Tinggi		1209 Simalungun		1272 Tanjung Balai
	1275 Medan		1212 Deli Serdang		1276 Binjai
	1277 Padang Sidempuan		1214 Nias Selatan		1404 Pelalawan
	1309 Pasaman		1225 Nias Barat		1471 Pekanbaru
	1310 Solok Selatan		1271 Sibolga		1473 Dumai
	1311 Dharmasraya		1278 Gunung Sitoli		1505 Muaro Jambi
	1371 Padang		1301 Kep. Mentawai		1602 Ogan Komering Ilir
	1373 Sawah Lunto		1302 Pesisir Selatan		1604 Lahat
	1375 Bukittinggi		1303 Solok		1605 Musi Rawas
	1377 Pariaman		1304 Sijunjung		1606 Musi Banyuasin
	1402 Indragiri Hulu		1305 Tanah Datar		1609 OKU Timur
	1405 Siak		1306 Padang Pariaman		1610 Ogan Ilir
	1406 Kampar		1307 Agam		1673 Pagar Alam
	1407 Rokan Hulu		1308 Limapuluh kota		1703 Bengkulu Utara
	1408 Bengkulu		1312 Pasaman Barat		1707 Lebong
	1410 Kep. Meranti		1372 Solok		1801 Lampung Barat
	1501 Kerinci		1374 Padang Panjang		1803 Lampung Selatan
	1502 Merangin		1376 Payakumbuh		1805 Lampung Tengah
	1503 Sarolangun		1401 Kuantan Singingi		1807 Way Kanan
	1504 Batang Hari		1403 Indragiri Hilir		1810 Pringsewu
	1507 Tanjung Jabung Barat		1409 Rokan Hilir		1812 Tulangbawang Barat
	1508 Tebo		1506 Tanjung Jabung Timur		1872 Metro
	1509 Bungo		1571 Jambi		2172 Tanjung Pinang
	1601 Ogan Komering Ulu				

1603	Muara Enim	1572	Sungai Penuh	3101	Kep. Seribu
1607	Banyu Asin	1608	OKU Selatan	3301	Cilacap
1672	Prabumulih	1611	Empat Lawang	3309	Boyolali
1702	Rejang Lebong	1671	Palembang	3310	Klaten
1706	Mukomuko	1674	Lubuk Linggau	3311	Sukoharjo
1771	Bengkulu	1701	Bengkulu Selatan	3314	Sragen
1802	Tanggamus	1704	Kaur	3316	Blora
1804	Lampung Timur	1705	Seluma	3324	Kendal
1808	Tulangbawang	1708	Kepahiang	3404	Sleman
1809	Pesawaran	1709	Bengkulu Tengah	3502	Ponorogo
1811	Mesuji	1806	Lampung Utara	3503	Trenggalek
1904	Bangka Tengah	1871	Bandar Lampung	3511	Bondowoso
2101	Karimun	1901	Bangka	3516	Mojokerto
2102	Bintan	1902	Belitung	3518	Nganjuk
2104	Lingga	1903	Bangka Barat	3571	Kediri
2171	Batam	1905	Bangka Selatan	3573	Malang
3174	Jakarta Barat	1906	Belitung Timur	3602	Lebak
3208	Kuningan	1971	Pangkal Pinang	3674	Tangerang Selatan
3209	Cirebon	2103	Natuna	5104	Gianyar
3212	Indramayu	2105	Kep. Anambas	5171	Denpasar
3215	Karawang	3171	Jakarta Selatan	5313	Manggarai
3216	Bekasi	3172	Jakarta Timur	6108	Kapuas Hulu
3273	Bandung	3173	Jakarta Pusat	6110	Melawi
3274	Cirebon	3175	Jakarta Utara	6171	Pontianak
3279	Banjar	3201	Bogor	6205	Barito Utara
3306	Purworejo	3202	Sukabumi	6210	Pulang Pisau
3313	Karanganyar	3203	Cianjur	6271	Palangkaraya
3315	Grobogan	3204	Bandung	6307	Hulu Sungai Tengah
3317	Rembang	3205	Garut	6371	Banjarmasin
3318	Pati	3206	Tasikmalaya	6407	Bulangan
3319	Kudus	3207	Ciamis	6409	Penajem Paser Utara
3320	Jepara	3210	Majalengka	6474	Bontang
3321	Demak	3211	Sumedang	7105	Minahasa Selatan
3326	Pekalongan	3213	Subang	7207	Buol
3375	Pekalongan	3214	Purwakarta	7314	Sidenreng Rappang
3504	Tulungagung	3217	Bandung Barat	7407	Wakatobi
3505	Blitar	3271	Bogor	7409	Buton Utara
3506	Kediri	3272	Sukabumi	7603	Mamasa
3514	Pasuruan	3275	Bekasi	8106	Seram Bagian Barat
3517	Jombang	3276	Depok	8171	Ambon
3519	Madiun	3277	Cimahi	9103	Teluk Wondama
3520	Magetan	3278	Tasikmalaya	9104	Teluk Bintuni
3522	Bojonegoro	3302	Banyumas	9105	Manokwari

3523	Tuban	3303	Purbalingga	9106	Sorong Selatan
3524	Lamongan	3304	Banjarnegara	9109	Tambrau
3526	Bangkalan	3305	Kebumen	9110	Maybrat
3572	Blitar	3307	Wonosobo	9401	Merauke
3577	Madiun	3308	Magelang	9403	Jayapura
3604	Serang	3312	Wonogiri	9409	Biak Numfor
3673	Serang	3322	Semarang	9415	Asmat
5101	Jembrana	3323	Temanggung	9419	Sarmi
5102	Tabanan	3325	Batang	9420	Keerom
5103	Badung	3327	Pemalang	9429	Nduga
5108	Buleleng	3328	Tegal	9431	Mamberamo Tengah
5201	Lombok Barat	3329	Brebes	9433	Puncak
5204	Sumbawa	3371	Magelang	9434	Dogiyai
5208	Lombok Utara	3372	Surakarta		
5304	TTS	3373	Salatiga		
5306	Belu	3374	Semarang		
5371	Kupang	3376	Tegal		
6101	Sambas	3401	Kulon Progo		
6103	Landak	3402	Bantul		
6105	Sanggau	3403	Gunung Kidul		
6106	Ketapang	3471	Yogyakarta		
6107	Sintang	3501	Pacitan		
6112	Kubu Raya	3507	Malang		
6172	Singkawang	3508	Lumajang		
6202	Kotawaringin Timur	3509	Jember		
6204	Barito Selatan	3510	Banyuwangi		
6206	Sukamara	3512	Situbondo		
6207	Lamandau	3513	Probolinggo		
6208	Seruyan	3515	Sidoarjo		
6212	Barito Timur	3521	Ngawi		
6213	Murung Raya	3525	Gresik		
6305	Tapin	3527	Sampang		
6306	Hulu Sungai Selatan	3528	Pamekasan		
6308	Hulu Sungai Utara	3529	Sumenep		
6309	Tabalong	3574	Probolinggo		
6372	Banjar Baru	3575	Pasuruan		
6401	Paser	3576	Mojokerto		
6403	Kutai Kartanegara	3578	Surabaya		
6404	Kutai Timur	3579	Batu		
6405	Berau	3601	Pandeglang		
6406	Malinau	3603	Tangerang		
6408	Tana Tidung	3671	Tangerang		
6410	Nunukan	3672	Cilegon		

6471	Balikpapan	5105	Klungkung		
6472	Samarinda	5106	Bangli		
7102	Minahasa	5107	Karang asem		
7103	Kep. Sangihe	5202	Lombok Tengah		
7171	Manado	5203	Lombok Timur		
7271	Palu	5205	Dompu		
7301	Kep. Selayar	5206	Bima		
7309	Pangkajene dan Kep.	5207	Sumbawa Barat		
7406	Bombana	5271	Mataram		
7408	Kolaka Utara	5272	Bima		
7410	Konawe Utara	5301	Sumba Barat		
7471	Kendari	5302	Sumba Timur		
7501	Boalemo	5303	Kupang		
7502	Gorontalo	5305	TTU		
7505	Gorontalo Utara	5307	Alor		
7601	Majene	5308	Lembata		
7604	Mamuju	5309	Flores Timur		
8101	Maluku Tenggara Barat	5310	Sikka		
8103	Maluku Tengah	5311	Ende		
8105	Kep. Aru	5312	Ngada		
8201	Halmahera Barat	5314	Rote Ndao		
8205	Halmahera Utara	5315	Manggarai Barat		
8272	Tidore Kepulauan	5316	Sumba Tengah		
9102	Kaimana	5317	Sumba Barat Daya		
9107	Sorong	5318	Nagekeo		
9108	Raja Ampat	5319	Manggarai Timur		
9404	Nabire	5320	Sabu Raijua		
9410	Panjai	6102	Bengkayang		
9413	Boven Digoel	6104	Mempawah		
9414	Mappi	6109	Sekadau		
9416	Yahukimo	6111	Kayong Utara		
9427	Supiori	6201	Kotawaringin Barat		
9436	Deiyai	6203	Kapuas		
9471	Jayapura	6209	Katingan		
		6211	Gunung Mas		
		6301	Tanah Laut		
		6302	Kota Baru		
		6303	Banjar		
		6304	Barito Kuala		
		6310	Tanah Bumbu		
		6311	Balangan		
		6402	Kutai Barat		
		6473	Tarakan		

		7101	Bolaang Mongondow		
		7104	Kep. Talaud		
		7106	Minahasa Utara		
		7107	Bolaang Mongondow Utara		
		7108	Siau Tagu Landang Biaro		
		7109	Minahasa Tenggara		
		7110	Bolaang Mongondow Selatan		
		7111	Bolaang Mongondow Timur		
		7172	Bitung		
		7173	Tomohon		
		7174	Kotamobagu		
		7201	Banggai Kepulauan		
		7202	Banggai		
		7203	Morowali		
		7204	Poso		
		7205	Donggala		
		7206	Toli-toli		
		7208	Parigi Moutong		
		7209	Tojo Una-una		
		7210	Sigi		
		7302	Bulukumba		
		7303	Bantaeng		
		7304	Jeneponto		
		7305	Takalar		
		7306	Gowa		
		7307	Sinjai		
		7308	Maros		
		7310	Barru		
		7311	Bone		
		7312	Soppeng		
		7313	Wajo		
		7315	Pinrang		
		7316	Enrekang		
		7317	Luwu		
		7318	Tana Toraja		
		7322	Luwu Utara		
		7325	Luwu Timur		
		7326	Toraja Utara		
		7371	Makassar		
		7372	Parepare		
		7373	Palopo		
		7401	Buton		

		7402	Muna		
		7403	Konawe		
		7404	Kolaka		
		7405	Konawe Selatan		
		7472	BauBau		
		7503	Pohuwato		
		7504	Bone Bolango		
		7571	Gorontalo		
		7602	Polewali Mandar		
		7605	Mamuju Utara		
		8102	Maluku Tenggara		
		8104	Buru		
		8107	Seram Bagian Timur		
		8108	Maluku Barat daya		
		8109	Buru Selatan		
		8172	Tual		
		8202	Halmahera Tengah		
		8203	Kep. Sula		
		8204	Halmahera Selatan		
		8206	Halmahera Timur		
		8207	Pulau Morotai		
		8271	Ternate		
		9101	Fakfak		
		9171	Sorong		
		9402	Jayawijaya		
		9408	Kep. Yapen		
		9411	Puncak Jaya		
		9412	Mimika		
		9417	Peg. Bintang		
		9418	Tolikara		
		9426	Waropen		
		9428	Mamberamo Raya		
		9430	Lanny Jaya		
		9432	Yalimo		
		9435	Intan Jaya		

**Lampiran 2f. Hasil Pengelompokan Kabupaten/Kota menurut Pola Makan
dengan Metode Kernel K-Means dengan k=3**

Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota	
1	1101	Simeulue	2	1102	Aceh Singkil	3	1105	Aceh Timur
	1103	Aceh Selatan		1104	Aceh Tenggara		1175	Subulussalam
	1106	Aceh Tengah		1110	Bireun		1203	TAPANULI SELATAN
	1107	Aceh Barat		1114	Aceh Tamiang		1218	Serdang Bedagai
	1108	Aceh Besar		1171	Banda Aceh		1220	Padang Lawas Utara
	1109	Pidie		1172	Sabang		1221	Padang Lawas
	1111	Aceh Utara		1173	Langsa		1271	Sibolga
	1112	Aceh Barat Daya		1174	Lhokseumawe		1302	Pesisir Selatan
	1113	Gayo Lues		1201	Nias		1303	Solok
	1115	Nagan Raya		1202	Mandailing Natal		1306	Padang Pariaman
	1116	Aceh Jaya		1204	TAPANULI TENGAH		1308	Limapuluh kota
	1117	Bener Meriah		1206	TOBA SAMOSIR		1371	Padang
	1118	Pidie Jaya		1207	LABUHAN BATU		1377	Pariaman
	1205	TAPANULI UTARA		1208	ASAHAN		1405	Siak
	1210	Dairi		1209	Simalungun		1408	Bengkalis
	1215	Humbang Hasundutan		1211	Karo		1505	Muaro Jambi
	1216	Pak Pak Bharat		1212	Deli Serdang		1572	Sungai Penuh
	1217	Samosir		1213	Langkat		1603	Muara Enim
	1225	Nias Barat		1214	Nias Selatan		1606	Musi Banyuasin
	1309	Pasaman		1219	Batu Bara		1671	Palembang
	1374	Padang Panjang		1222	Labuan Batu Selatan		1672	Prabumulih
	1375	Bukittinggi		1223	Labuan Batu Utara		1706	Mukomuko
	1403	Indragiri Hilir		1224	Nias Utara		1805	Lampung Tengah
	1410	Kep. Meranti		1272	Tanjung Balai		1901	Bangka
	1501	Kerinci		1273	Pematang Siantar		1902	Belitung
	1506	Tanjung Jabung Timur		1274	Tebing Tinggi		2105	Kep. Anambas
	1507	Tanjung Jabung Barat		1275	Medan		2171	Batam
	1601	Ogan Komering Ulu		1276	Binjai		3171	Jakarta Selatan
	1602	Ogan Komering Ilir		1277	Padang Sidempuan		3173	Jakarta Pusat
	1604	Lahat		1278	Gunung Sitoli		3174	Jakarta Barat
	1605	Musi Rawas		1301	Kep. Mentawai		3175	Jakarta Utara
	1607	Banyu Asin		1304	Sijunjung		3202	Sukabumi
	1608	OKU Selatan		1305	Tanah Datar		3205	Garut
	1609	OKU Timur		1307	Agam		3207	Ciamis
	1610	Ogan Ilir		1310	Solok Selatan		3212	Indramayu
	1611	Empat Lawang		1311	Dharmasraya		3214	Purwakarta
	1673	Pagar Alam		1312	Pasaman Barat		3274	Cirebon
				1372	Solok		3275	Bekasi

1674	Lubuk Linggau	1373	Sawah Lunto	3276	Depok
1701	Bengkulu Selatan	1376	Payakumbuh	3302	Banyumas
1702	Rejang Lebong	1401	Kuantan Singingi	3303	Purbalingga
1703	Bengkulu Utara	1402	Indragiri Hulu	3306	Purworejo
1704	Kaur	1404	Pelalawan	3315	Grobogan
1707	Lebong	1406	Kampar	3376	Tegal
1708	Kepahiang	1407	Rokan Hulu	3404	Sleman
1709	Bengkulu Tengah	1409	Rokan Hilir	3510	Banyuwangi
1801	Lampung Barat	1471	Pekanbaru	3514	Pasuruan
1802	Tanggamus	1473	Dumai	3515	Sidoarjo
1803	Lampung Selatan	1502	Merangin	3518	Nganjuk
1804	Lampung Timur	1503	Sarolangun	3524	Lamongan
1806	Lampung Utara	1504	Batang Hari	3526	Bangkalan
1807	Way Kanan	1508	Tebo	3529	Sumenep
1808	Tulangbawang	1509	Bungo	3571	Kediri
1809	Pesawaran	1571	Jambi	3572	Blitar
1810	Pringsewu	1705	Seluma	3602	Lebak
1811	Mesuji	1771	Bengkulu	3671	Tangerang
1812	Tulangbawang Barat	1872	Metro	3672	Cilegon
1871	Bandar Lampung	1906	Belitung Timur	3673	Serang
1903	Bangka Barat	1971	Pangkal Pinang	5171	Denpasar
1904	Bangka Tengah	2101	Karimun	5305	TTU
1905	Bangka Selatan	2102	Bintan	5371	Kupang
3101	Kep. Seribu	2103	Natuna	6105	Sanggau
3172	Jakarta Timur	2104	Lingga	6112	Kubu Raya
3201	Bogor	2172	Tanjung Pinang	6171	Pontianak
3203	Cianjur	3324	Kendal	6207	Lamandau
3204	Bandung	3528	Pamekasan	6208	Seruyan
3206	Tasikmalaya	3674	Tangerang Selatan	6210	Pulang Pisau
3208	Kuningan	5104	Gianyar	6211	Gunung Mas
3209	Cirebon	5204	Sumbawa	6301	Tanah Laut
3210	Majalengka	5205	Dompu	6302	Kota Baru
3211	Sumedang	5206	Bima	6307	Hulu Sungai Tengah
3213	Subang	5207	Sumbawa Barat	6409	Penajem Paser Utara
3215	Karawang	5272	Bima	6471	Balikpapan
3216	Bekasi	5304	TTS	7101	Bolaang Mongondow
3217	Bandung Barat	5306	Belu	7102	Minahasa
3271	Bogor	5308	Lembata	7106	Minahasa Utara
3272	Sukabumi	5314	Rote Ndao	7171	Manado
3273	Bandung	5320	Sabu Raijua	7172	Bitung
3277	Cimahi	6206	Sukamara	7173	Tomohon
3278	Tasikmalaya	6306	Hulu Sungai Selatan	7174	Kotamobagu
3279	Banjar	6310	Tanah Bumbu	7202	Banggai

3301	Cilacap	6311	Balangan	7203	Morowali
3304	Banjarnegara	6372	Banjar Baru	7206	Toli-toli
3305	Kebumen	6401	Paser	7208	Parigi Moutong
3307	Wonosobo	6403	Kutai Kartanegara	7209	Tojo Una-una
3308	Magelang	6404	Kutai Timur	7210	Sigi
3309	Boyolali	6405	Berau	7302	Bulukumba
3310	Klaten	6410	Nunukan	7305	Takalar
3311	Sukoharjo	6472	Samarinda	7312	Soppeng
3312	Wonogiri	6473	Tarakan	7313	Wajo
3313	Karanganyar	6474	Bontang	7314	Sidenreng Rappang
3314	Sragen	7104	Kep. Talaud	7315	Pinrang
3316	Blora	7201	Banggai Kepulauan	7317	Luwu
3317	Rembang	7205	Donggala	7322	Luwu Utara
3318	Pati	7271	Palu	7325	Luwu Timur
3319	Kudus	7301	Kep. Selayar	7371	Makassar
3320	Jepara	7304	Jeneponto	7373	Palopo
3321	Demak	7307	Sinjai	7403	Konawe
3322	Semarang	7308	Maros	7406	Bombana
3323	Temanggung	7309	Pangkajene dan Kep.	7501	Boalemo
3325	Batang	7310	Barru	7502	Gorontalo
3326	Pekalongan	7311	Bone	7503	Pohuwato
3327	Pemalang	7372	Parepare	7504	Bone Bolango
3328	Tegal	7401	Buton	7601	Majene
3329	Brebes	7402	Muna	7602	Polewali Mandar
3371	Magelang	7404	Kolaka	7605	Mamuju Utara
3372	Surakarta	7405	Konawe Selatan	8103	Maluku Tengah
3373	Salatiga	7407	Wakatobi	8104	Buru
3374	Semarang	7408	Kolaka Utara	8106	Seram Bagian Barat
3375	Pekalongan	7409	Buton Utara	8107	Seram Bagian Timur
3401	Kulon Progo	7410	Konawe Utara	8172	Tual
3402	Bantul	7471	Kendari	8202	Halmahera Tengah
3403	Gunung Kidul	7472	BauBau	8203	Kep. Sula
3471	Yogyakarta	8101	Maluku Tenggara Barat	8204	Halmahera Selatan
3501	Pacitan	8102	Maluku Tenggara	8206	Halmahera Timur
3502	Ponorogo	8108	Maluku Barat daya	9101	Fakfak
3503	Trenggalek	8109	Buru Selatan	9106	Sorong Selatan
3504	Tulungagung	8201	Halmahera Barat	9107	Sorong
3505	Blitar	8205	Halmahera Utara	9401	Merauke
3506	Kediri	8272	Tidore Kepulauan	9402	Jayawijaya
3507	Malang	9102	Kaimana	9404	Nabire
3508	Lumajang	9103	Teluk Wondama	9409	Biak Numfor
3509	Jember	9104	Teluk Bintuni	9412	Mimika
3511	Bondowoso	9105	Manokwari	9432	Yalimo

3512	Situbondo	9108	Raja Ampat	9471	Jayapura
3513	Probolinggo	9109	Tambrau		
3516	Mojokerto	9171	Sorong		
3517	Jombang	9403	Jayapura		
3519	Madiun	9408	Kep. Yape		
3520	Magetan	9410	Panai		
3521	Ngawi	9411	Puncak Jaya		
3522	Bojonegoro	9413	Boven Digoel		
3523	Tuban	9414	Mappi		
3525	Gresik	9416	Yahukimo		
3527	Sampang	9417	Peg. Bintang		
3573	Malang	9418	Tolikara		
3574	Probolinggo	9419	Sarmi		
3575	Pasuruan	9420	Keerom		
3576	Mojokerto	9428	Mamberamo Raya		
3577	Madiun	9429	Nduga		
3578	Surabaya	9431	Mamberamo Tengah		
3579	Batu	9433	Puncak		
3601	Pandeglang				
3603	Tangerang				
3604	Serang				
5101	Jembrana				
5102	Tabanan				
5103	Badung				
5105	Klungkung				
5106	Bangli				
5107	Karang asem				
5108	Buleleng				
5201	Lombok Barat				
5202	Lombok Tengah				
5203	Lombok Timur				
5208	Lombok Utara				
5271	Mataram				
5301	Sumba Barat				
5302	Sumba Timur				
5303	Kupang				
5307	Alor				
5309	Flores Timur				
5310	Sikka				
5311	Ende				
5312	Ngada				
5313	Manggarai				
5315	Manggarai Barat				

5316	Sumba Tengah				
5317	Sumba Barat Daya				
5318	Nagekeo				
5319	Manggarai Timur				
6101	Sambas				
6102	Bengkayang				
6103	Landak				
6104	Mempawah				
6106	Ketapang				
6107	Sintang				
6108	Kapuas Hulu				
6109	Sekadau				
6110	Melawi				
6111	Kayong Utara				
6172	Singkawang				
6201	Kotawaringin Barat				
6202	Kotawaringin Timur				
6203	Kapuas				
6204	Barito Selatan				
6205	Barito Utara				
6209	Katingan				
6212	Barito Timur				
6213	Murung Raya				
6271	Palangkaraya				
6303	Banjar				
6304	Barito Kuala				
6305	Tapin				
6308	Hulu Sungai Utara				
6309	Tabalong				
6371	Banjarmasin				
6402	Kutai Barat				
6406	Malinau				
6407	Bulungan				
6408	Tana Tidung				
7103	Kep. Sangihe				
7105	Minahasa Selatan				
7107	Bolaang Mongondow Utara				
7108	Siau Tagu Landang Biaro				
7109	Minahasa Tenggara				
7110	Bolaang Mongondow Selatan				
7111	Bolaang Mongondow Timur				
7204	Poso				

7207	Buol				
7303	Bantaeng				
7306	Gowa				
7316	Enrekang				
7318	Tana Toraja				
7326	Toraja Utara				
7505	Gorontalo Utara				
7571	Gorontalo				
7603	Mamasa				
7604	Mamuju				
8105	Kep. Aru				
8171	Ambon				
8207	Pulau Morotai				
8271	Ternate				
9110	Maybrat				
9415	Asmat				
9426	Waropen				
9427	Supiori				
9430	Lanny Jaya				
9434	Dogiyai				
9435	Intan Jaya				
9436	Deiyai				

**Lampiran 2g. Hasil Pengelompokan Kabupaten/Kota menurut Aktivitas Fisik
dengan Metode Kernel K-Means dengan k=3**

Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota	
1	1103	Aceh Selatan	2	1102	Aceh Singkil	3	1101	Simeulue
	1108	Aceh Besar		1105	Aceh Timur		1104	Aceh Tenggara
	1109	Pidie		1106	Aceh Tengah		1111	Aceh Utara
	1171	Banda Aceh		1107	Aceh Barat		1115	Nagan Raya
	1173	Langsa		1110	Bireun		1116	Aceh Jaya
	1174	Lhokseumawe		1112	Aceh Barat Daya		1117	Bener Meriah
	1207	LABUHAN BATU		1113	Gayo Lues		1118	Pidie Jaya
	1218	Serdang Bedagai		1114	Aceh Tamiang		1201	Nias
	1222	Labuan Batu Selatan		1172	Sabang		1202	Mandailing Natal
	1271	Sibolga		1175	Subulussalam		1203	TAPANULI SELATAN
	1272	Tanjung Balai		1208	ASAHAN		1204	TAPANULI TENGAH
	1274	Tebing Tinggi		1209	Simalungun		1205	TAPANULI UTARA
	1275	Medan		1212	Deli Serdang		1206	TOBA SAMOSIR
	1276	Binjai		1213	Langkat		1210	Dairi
	1302	Pesisir Selatan		1219	Batu Bara		1211	Karo
	1371	Padang		1221	Padang Lawas		1214	Nias Selatan
	1372	Solok		1223	Labuan Batu Utara		1215	Humbang Hasundutan
	1373	Sawah Lunto		1273	Pematang Siantar		1216	Pak Pak Bharat
	1374	Padang Panjang		1277	Padang Sidempuan		1217	Samosir
	1375	Bukittinggi		1278	Gunung Sitoli		1220	Padang Lawas Utara
	1376	Payakumbuh		1303	Solok		1224	Nias Utara
	1377	Pariaman		1304	Sijunjung		1225	Nias Barat
	1404	Pelalawan		1305	Tanah Datar		1301	Kep. Mentawai
	1405	Siak		1306	Padang Pariaman		1309	Pasaman
	1408	Bengkalis		1307	Agam		1310	Solok Selatan
	1409	Rokan Hilir		1308	Limapuluh kota		1407	Rokan Hulu
	1410	Kep. Meranti		1311	Dharmasraya		1501	Kerinci
	1471	Pekanbaru		1312	Pasaman Barat		1502	Merangin
	1504	Batang Hari		1401	Kuantan Singingi		1505	Muaro Jambi
	1571	Jambi		1402	Indragiri Hulu		1506	Tanjung Jabung Timur
	1671	Palembang		1403	Indragiri Hilir		1603	Muara Enim
	1706	Mukomuko		1406	Kampar		1604	Lahat
	1771	Bengkulu		1473	Dumai		1607	Banyu Asin
	1802	Tanggamus		1503	Sarolangun		1608	OKU Selatan
	1810	Pringsewu		1507	Tanjung Jabung Barat		1609	OKU Timur
	1811	Mesuji		1508	Tebo		1611	Empat Lawang
	1871	Bandar Lampung		1509	Bungo		1674	Lubuk Linggau
	1872	Metro		1572	Sungai Penuh		1703	Bengkulu Utara
	1904	Bangka Tengah						

1971	Pangkal Pinang	1601	Ogan Komering Ulu	1704	Kaur
2101	Karimun	1602	Ogan Komering Ilir	1705	Seluma
2102	Bintan	1605	Musi Rawas	1707	Lebong
2103	Natuna	1606	Musi Banyuasin	1709	Bengkulu Tengah
2105	Kep. Anambas	1610	Ogan Ilir	1801	Lampung Barat
2172	Tanjung Pinang	1672	Prabumulih	1804	Lampung Timur
3171	Jakarta Selatan	1673	Pagar Alam	1812	Tulangbawang Barat
3172	Jakarta Timur	1701	Bengkulu Selatan	1903	Bangka Barat
3173	Jakarta Pusat	1702	Rejang Lebong	1905	Bangka Selatan
3174	Jakarta Barat	1708	Kepahiang	2104	Lingga
3175	Jakarta Utara	1803	Lampung Selatan	3202	Sukabumi
3203	Cianjur	1805	Lampung Tengah	3215	Karawang
3204	Bandung	1806	Lampung Utara	3305	Kebumen
3208	Kuningan	1807	Way Kanan	3308	Magelang
3210	Majalengka	1808	Tulangbawang	3312	Wonogiri
3214	Purwakarta	1809	Pesawaran	3313	Karanganyar
3216	Bekasi	1901	Bangka	3315	Grobogan
3271	Bogor	1902	Belitung	3316	Blora
3272	Sukabumi	1906	Belitung Timur	3323	Temanggung
3273	Bandung	2171	Batam	3329	Brebes
3274	Cirebon	3101	Kep. Seribu	3375	Pekalongan
3275	Bekasi	3201	Bogor	3403	Gunung Kidul
3276	Depok	3205	Garut	3501	Pacitan
3277	Cimahi	3206	Tasikmalaya	3508	Lumajang
3278	Tasikmalaya	3207	Ciamis	3511	Bondowoso
3279	Banjar	3209	Cirebon	3522	Bojonegoro
3303	Purbalingga	3211	Sumedang	3523	Tuban
3311	Sukoharjo	3212	Indramayu	3527	Sampang
3319	Kudus	3213	Subang	5105	Klungkung
3320	Jepara	3217	Bandung Barat	5106	Bangli
3328	Tegal	3301	Cilacap	5204	Sumbawa
3371	Magelang	3302	Banyumas	5207	Sumbawa Barat
3372	Surakarta	3304	Banjarnegara	5301	Sumba Barat
3373	Salatiga	3306	Purworejo	5302	Sumba Timur
3374	Semarang	3307	Wonosobo	5303	Kupang
3376	Tegal	3309	Boyolali	5304	TTS
3402	Bantul	3310	Klaten	5305	TTU
3404	Sleman	3314	Sragen	5306	Belu
3471	Yogyakarta	3317	Rembang	5308	Lembata
3515	Sidoarjo	3318	Pati	5309	Flores Timur
3516	Mojokerto	3321	Demak	5311	Ende
3517	Jombang	3322	Semarang	5312	Ngada
3525	Gresik	3324	Kendal	5314	Rote Ndao

3571	Kediri	3325	Batang	5315	Manggarai Barat
3572	Blitar	3326	Pekalongan	5316	Sumba Tengah
3573	Malang	3327	Pemalang	5317	Sumba Barat Daya
3574	Probolinggo	3401	Kulon Progo	5318	Nagekeo
3575	Pasuruan	3502	Ponorogo	5319	Manggarai Timur
3577	Madiun	3503	Trenggalek	6101	Sambas
3578	Surabaya	3504	Tulungagung	6102	Bengkayang
3579	Batu	3505	Blitar	6103	Landak
3602	Lebak	3506	Kediri	6105	Sanggau
3603	Tangerang	3507	Malang	6106	Ketapang
3671	Tangerang	3509	Jember	6109	Sekadau
3672	Cilegon	3510	Banyuwangi	6111	Kayong Utara
3673	Serang	3512	Situbondo	6205	Barito Utara
3674	Tangerang Selatan	3513	Probolinggo	6211	Gunung Mas
5102	Tabanan	3514	Pasuruan	6213	Murung Raya
5103	Badung	3518	Nganjuk	6372	Banjar Baru
5108	Buleleng	3519	Madiun	6410	Nunukan
5171	Denpasar	3520	Magetan	6471	Balikpapan
5205	Dompu	3521	Ngawi	7104	Kep. Talaud
5271	Mataram	3524	Lamongan	7107	Bolaang Mongondow Utara
5272	Bima	3526	Bangkalan	7110	Bolaang Mongondow Selatan
5371	Kupang	3528	Pamekasan	7201	Banggai Kepulauan
6112	Kubu Raya	3529	Sumenep	7203	Morowali
6171	Pontianak	3576	Mojokerto	7204	Poso
6210	Pulang Pisau	3601	Pandeglang	7206	Toli-toli
6212	Barito Timur	3604	Serang	7207	Buol
6271	Palangkaraya	5101	Jembrana	7208	Parigi Moutong
6306	Hulu Sungai Selatan	5104	Gianyar	7210	Sigi
6308	Hulu Sungai Utara	5107	Karang asem	7302	Bulukumba
6310	Tanah Bumbu	5201	Lombok Barat	7303	Bantaeng
6311	Balangan	5202	Lombok Tengah	7306	Gowa
6371	Banjarmasin	5203	Lombok Timur	7307	Sinjai
6401	Paser	5206	Bima	7316	Enrekang
6405	Berau	5208	Lombok Utara	7317	Luwu
6472	Samarinda	5307	Alor	7318	Tana Toraja
6473	Tarakan	5310	Sikka	7322	Luwu Utara
7171	Manado	5313	Manggarai	7325	Luwu Timur
7172	Bitung	5320	Sabu Raijua	7326	Toraja Utara
7174	Kotamobagu	6104	Mempawah	7401	Buton
7202	Banggai	6107	Sintang	7402	Muna
7271	Palu	6108	Kapuas Hulu	7404	Kolaka
7301	Kep. Selayar	6110	Melawi	7405	Konawe Selatan
7310	Barru	6172	Singkawang	7406	Bombana

7371	Makassar	6201	Kotawaringin Barat	7408	Kolaka Utara
7372	Parepare	6202	Kotawaringin Timur	7503	Pohuwato
7373	Palopo	6203	Kapuas	7505	Gorontalo Utara
7403	Konawe	6204	Barito Selatan	7601	Majene
7407	Wakatobi	6206	Sukamara	7604	Mamuju
7471	Kendari	6207	Lamandau	7605	Mamuju Utara
7472	BauBau	6208	Seruyan	8101	Maluku Tenggara Barat
7504	Bone Bolango	6209	Katingan	8102	Maluku Tenggara
7571	Gorontalo	6301	Tanah Laut	8105	Kep. Aru
8104	Buru	6302	Kota Baru	8106	Seram Bagian Barat
8171	Ambon	6303	Banjar	8107	Seram Bagian Timur
8172	Tual	6304	Barito Kuala	8108	Maluku Barat daya
8271	Ternate	6305	Tapin	8109	Buru Selatan
8272	Tidore Kepulauan	6307	Hulu Sungai Tengah	8202	Halmahera Tengah
9101	Fakfak	6309	Tabalong	8203	Kep. Sula
9105	Manokwari	6402	Kutai Barat	8204	Halmahera Selatan
9171	Sorong	6403	Kutai Kartanegara	8205	Halmahera Utara
9471	Jayapura	6404	Kutai Timur	8206	Halmahera Timur
		6406	Malinau	8207	Pulau Morotai
		6407	Bulungan	9102	Kaimana
		6408	Tana Tidung	9103	Teluk Wondama
		6409	Penajem Paser Utara	9104	Teluk Bintuni
		6474	Bontang	9106	Sorong Selatan
		7101	Bolaang Mongondow	9107	Sorong
		7102	Minahasa	9109	Tambrauw
		7103	Kep. Sangihe	9110	Maybrat
		7105	Minahasa Selatan	9402	Jayawijaya
		7106	Minahasa Utara	9403	Jayapura
		7108	Siau Tagu Landang Biaro	9409	Biak Numfor
		7109	Minahasa Tenggara	9410	Panjai
		7111	Bolaang Mongondow Timur	9411	Puncak Jaya
		7173	Tomohon	9413	Boven Digoel
		7205	Donggala	9414	Mappi
		7209	Tojo Una-una	9416	Yahukimo
		7304	Jeneponto	9417	Peg. Bintang
		7305	Takalar	9418	Tolikara
		7308	Maros	9427	Supiori
		7309	Pangkajene dan Kep.	9428	Mamberamo Raya
		7311	Bone	9429	Nduga
		7312	Soppeng	9430	Lanny Jaya
		7313	Wajo	9431	Mamberamo Tengah
		7314	Sidenreng Rappang	9432	Yalimo
		7315	Pinrang	9433	Puncak

		7409	Buton Utara	9434	Dogiyai
		7410	Konawe Utara	9435	Intan Jaya
		7501	Boalemo	9436	Deiyai
		7502	Gorontalo		
		7602	Polewali Mandar		
		7603	Mamasa		
		8103	Maluku Tengah		
		8201	Halmahera Barat		
		9108	Raja Ampat		
		9401	Merauke		
		9404	Nabire		
		9408	Kep. Yapen		
		9412	Mimika		
		9415	Asmat		
		9419	Sarmi		
		9420	Keerom		
		9426	Waropen		

**Lampiran 2h. Hasil Pengelompokan Kabupaten/Kota menurut Gaya Hidup
dengan Metode Kernel K-Means dengan k=3**

Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota	
1	1101	Simeulue	2	1109	Pidie	3	1105	Aceh Timur
	1102	Aceh Singkil		1112	Aceh Barat Daya		1106	Aceh Tengah
	1103	Aceh Selatan		1115	Nagan Raya		1113	Gayo Lues
	1104	Aceh Tenggara		1201	Nias		1116	Aceh Jaya
	1107	Aceh Barat		1214	Nias Selatan		1117	Bener Meriah
	1108	Aceh Besar		1223	Labuan Batu Utara		1172	Sabang
	1110	Bireun		1224	Nias Utara		1175	Subulussalam
	1111	Aceh Utara		1225	Nias Barat		1202	Mandailing Natal
	1114	Aceh Tamiang		1278	Gunung Sitoli		1203	TAPANULI SELATAN
	1118	Pidie Jaya		3311	Sukoharjo		1204	TAPANULI TENGAH
	1171	Banda Aceh		3313	Karanganyar		1209	Simalungun
	1173	Langsa		3314	Sragen		1210	Dairi
	1174	Lhokseumawe		3315	Grobogan		1211	Karo
	1205	TAPANULI UTARA		3317	Rembang		1212	Deli Serdang
	1206	TOBA SAMOSIR		3318	Pati		1216	Pak Pak Bharat
	1207	LABUHAN BATU		3375	Pekalongan		1217	Samosir
	1208	ASAHAN		3376	Tegal		1220	Padang Lawas Utara
	1213	Langkat		3518	Nganjuk		1221	Padang Lawas
	1215	Humbang Hasundutan		3525	Gresik		1274	Tebing Tinggi
	1218	Serdang Bedagai		3571	Kediri		1277	Padang Sidempuan
	1219	Batu Bara		3575	Pasuruan		1301	Kep. Mentawai
	1222	Labuan Batu Selatan		5104	Gianyar		1302	Pesisir Selatan
	1271	Sibolga		5107	Karang asem		1303	Solok
	1272	Tanjung Balai		5171	Denpasar		1304	Sijunjung
	1273	Pematang Siantar		5303	Kupang		1305	Tanah Datar
	1275	Medan		5304	TTS		1307	Agam
	1276	Binjai		5306	Belu		1308	Limapuluh kota
	1306	Padang Pariaman		5308	Lembata		1309	Pasaman
	1311	Dharmasraya		5312	Ngada		1310	Solok Selatan
	1371	Padang		5317	Sumba Barat Daya		1312	Pasaman Barat
	1377	Pariaman		5320	Sabu Raijua		1372	Solok
	1401	Kuantan Singingi		5371	Kupang		1373	Sawah Lunto
	1403	Indragiri Hilir		6101	Sambas		1374	Padang Panjang
	1404	Pelalawan		6108	Kapuas Hulu		1375	Bukittinggi
	1405	Siak		6112	Kubu Raya		1376	Payakumbuh
	1406	Kampar		6171	Pontianak		1402	Indragiri Hulu
	1407	Rokan Hulu		6172	Singkawang		1507	Tanjung Jabung Barat
	1408	Bengkalis		6202	Kotawaringin Timur		1603	Muara Enim
	1409	Rokan Hilir		6203	Kapuas			

1410	Kep. Meranti	6210	Pulang Pisau	1605	Musi Rawas
1471	Pekanbaru	6271	Palangkaraya	1607	Banyu Asin
1473	Dumai	6305	Tapin	1608	OKU Selatan
1501	Kerinci	6306	Hulu Sungai Selatan	1673	Pagar Alam
1502	Merangin	6308	Hulu Sungai Utara	1674	Lubuk Linggau
1503	Sarolangun	6309	Tabalong	1702	Rejang Lebong
1504	Batang Hari	6371	Banjarmasin	1705	Seluma
1505	Muaro Jambi	6372	Banjar Baru	1707	Lebong
1506	Tanjung Jabung Timur	6404	Kutai Timur	1708	Kepahiang
1508	Tebo	6409	Penajem Paser Utara	1709	Bengkulu Tengah
1509	Bungo	6410	Nunukan	1801	Lampung Barat
1571	Jambi	7308	Maros	1802	Tanggamus
1572	Sungai Penuh	7309	Pangkajene dan Kep.	1803	Lampung Selatan
1601	Ogan Komering Ulu	7372	Parepare	1804	Lampung Timur
1602	Ogan Komering Ilir	7401	Buton	1807	Way Kanan
1604	Lahat	7402	Muna	1809	Pesawaran
1606	Musi Banyuasin	7407	Wakatobi	1811	Mesuji
1609	OKU Timur	7410	Konawe Utara	1871	Bandar Lampung
1610	Ogan Ilir	7471	Kendari	1903	Bangka Barat
1611	Empat Lawang	7472	BauBau	1905	Bangka Selatan
1671	Palembang	7601	Majene	1971	Pangkal Pinang
1672	Prabumulih	8106	Seram Bagian Barat	2101	Karimun
1701	Bengkulu Selatan	9103	Teluk Wondama	2105	Kep. Anambas
1703	Bengkulu Utara	9110	Maybrat	3201	Bogor
1704	Kaur	9171	Sorong	3202	Sukabumi
1706	Mukomuko	9402	Jayawijaya	3203	Cianjur
1771	Bengkulu	9403	Jayapura	3204	Bandung
1805	Lampung Tengah	9408	Kep. Yapen	3205	Garut
1806	Lampung Utara	9409	Biak Numfor	3206	Tasikmalaya
1808	Tulangbawang	9410	Panjai	3207	Ciamis
1810	Pringsewu	9411	Puncak Jaya	3208	Kuningan
1812	Tulangbawang Barat	9416	Yahukimo	3210	Majalengka
1872	Metro	9417	Peg. Bintang	3211	Sumedang
1901	Bangka	9418	Tolikara	3212	Indramayu
1902	Belitung	9419	Sarmi	3213	Subang
1904	Bangka Tengah	9426	Waropen	3214	Purwakarta
1906	Belitung Timur	9429	Nduga	3217	Bandung Barat
2102	Bintan	9430	Lanny Jaya	3271	Bogor
2103	Natuna	9431	Mamberamo Tengah	3272	Sukabumi
2104	Lingga	9433	Puncak	3275	Bekasi
2171	Batam	9434	Dogiyai	3277	Cimahi
2172	Tanjung Pinang	9435	Intan Jaya	3278	Tasikmalaya
3101	Kep. Seribu	9436	Deiyai	3279	Banjar

3171	Jakarta Selatan	9471	Jayapura	3302	Banyumas
3172	Jakarta Timur			3303	Purbalingga
3173	Jakarta Pusat			3304	Banjarnegara
3174	Jakarta Barat			3305	Kebumen
3175	Jakarta Utara			3307	Wonosobo
3209	Cirebon			3308	Magelang
3215	Karawang			3323	Temanggung
3216	Bekasi			3327	Pemalang
3273	Bandung			3371	Magelang
3274	Cirebon			3401	Kulon Progo
3276	Depok			3404	Sleman
3301	Cilacap			3471	Yogyakarta
3306	Purworejo			3502	Ponorogo
3309	Boyolali			3505	Blitar
3310	Klaten			3507	Malang
3312	Wonogiri			3508	Lumajang
3316	Blora			3509	Jember
3319	Kudus			3510	Banyuwangi
3320	Jepara			3512	Situbondo
3321	Demak			3513	Probolinggo
3322	Semarang			3529	Sumenep
3324	Kendal			3579	Batu
3325	Batang			3601	Pandeglang
3326	Pekalongan			3602	Lebak
3328	Tegal			3603	Tangerang
3329	Brebes			3604	Serang
3372	Surakarta			3672	Cilegon
3373	Salatiga			3673	Serang
3374	Semarang			5204	Sumbawa
3402	Bantul			5205	Dompu
3403	Gunung Kidul			5206	Bima
3501	Pacitan			5208	Lombok Utara
3503	Trenggalek			5302	Sumba Timur
3504	Tulungagung			5316	Sumba Tengah
3506	Kediri			5319	Manggarai Timur
3511	Bondowoso			6105	Sanggau
3514	Pasuruan			6107	Sintang
3515	Sidoarjo			6109	Sekadau
3516	Mojokerto			6110	Melawi
3517	Jombang			6111	Kayong Utara
3519	Madiun			6206	Sukamara
3520	Magetan			6207	Lamandau
3521	Ngawi			6208	Seruyan

3522	Bojonegoro			6209	Katingan
3523	Tuban			6211	Gunung Mas
3524	Lamongan			6310	Tanah Bumbu
3526	Bangkalan			6402	Kutai Barat
3527	Sampang			7101	Bolaang Mongondow
3528	Pamekasan			7102	Minahasa
3572	Blitar			7104	Kep. Talaud
3573	Malang			7105	Minahasa Selatan
3574	Probolinggo			7106	Minahasa Utara
3576	Mojokerto			7107	Bolaang Mongondow Utara
3577	Madiun			7108	Siau Tagu Landang Biaro
3578	Surabaya			7109	Minahasa Tenggara
3671	Tangerang			7110	Bolaang Mongondow Selatan
3674	Tangerang Selatan			7111	Bolaang Mongondow Timur
5101	Jembrana			7172	Bitung
5102	Tabanan			7173	Tomohon
5103	Badung			7174	Kotamobagu
5105	Klungkung			7201	Banggai Kepulauan
5106	Bangli			7203	Morowali
5108	Buleleng			7204	Poso
5201	Lombok Barat			7205	Donggala
5202	Lombok Tengah			7206	Toli-toli
5203	Lombok Timur			7207	Buol
5207	Sumbawa Barat			7209	Tojo Una-una
5271	Mataram			7210	Sigi
5272	Bima			7322	Luwu Utara
5301	Sumba Barat			7373	Palopo
5305	TTU			7408	Kolaka Utara
5307	Alor			7501	Boalemo
5309	Flores Timur			7502	Gorontalo
5310	Sikka			7503	Pohuwato
5311	Ende			7504	Bone Bolango
5313	Manggarai			7505	Gorontalo Utara
5314	Rote Ndao			7605	Mamuju Utara
5315	Manggarai Barat			8102	Maluku Tenggara
5318	Nagekeo			8105	Kep. Aru
6102	Bengkayang			8107	Seram Bagian Timur
6103	Landak			8201	Halmahera Barat
6104	Mempawah			8202	Halmahera Tengah
6106	Ketapang			8203	Kep. Sula
6201	Kotawaringin Barat			8204	Halmahera Selatan
6204	Barito Selatan			8206	Halmahera Timur
6205	Barito Utara				

6212	Barito Timur			8207	Pulau Morotai
6213	Murung Raya			8271	Ternate
6301	Tanah Laut			9102	Kaimana
6302	Kota Baru			9106	Sorong Selatan
6303	Banjar			9109	Tambrau
6304	Barito Kuala			9412	Mimika
6307	Hulu Sungai Tengah			9413	Boven Digoel
6311	Balangan			9414	Mappi
6401	Paser			9415	Asmat
6403	Kutai Kartanegara			9420	Keerom
6405	Berau			9428	Mamberamo Raya
6406	Malinau				
6407	Bulungan				
6408	Tana Tidung				
6471	Balikpapan				
6472	Samarinda				
6473	Tarakan				
6474	Bontang				
7103	Kep. Sangihe				
7171	Manado				
7202	Banggai				
7208	Parigi Moutong				
7271	Palu				
7301	Kep. Selayar				
7302	Bulukumba				
7303	Bantaeng				
7304	Jeneponto				
7305	Takalar				
7306	Gowa				
7307	Sinjai				
7310	Barru				
7311	Bone				
7312	Soppeng				
7313	Wajo				
7314	Sidenreng Rappang				
7315	Pinrang				
7316	Enrekang				
7317	Luwu				
7318	Tana Toraja				
7325	Luwu Timur				
7326	Toraja Utara				
7371	Makassar				
7403	Konawe				

7404	Kolaka				
7405	Konawe Selatan				
7406	Bombana				
7409	Buton Utara				
7571	Gorontalo				
7602	Polewali Mandar				
7603	Mamasa				
7604	Mamuju				
8101	Maluku Tenggara Barat				
8103	Maluku Tengah				
8104	Buru				
8108	Maluku Barat daya				
8109	Buru Selatan				
8171	Ambon				
8172	Tual				
8205	Halmahera Utara				
8272	Tidore Kepulauan				
9101	Fakfak				
9104	Teluk Bintuni				
9105	Manokwari				
9107	Sorong				
9108	Raja Ampat				
9401	Merauke				
9404	Nabire				
9427	Supiori				
9432	Yalimo				

**Lampiran 2i. Hasil Pengelompokan Kabupaten/Kota menurut Obesitas
dengan Metode Kernel K-Means dengan k=3**

Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota		Kel	Kode dan Nama Kab/Kota	
1	1102	Aceh Singkil	2	1103	Aceh Selatan	3	1101	Simeulue
	1104	Aceh Tenggara		1105	Aceh Timur		1115	Nagan Raya
	1106	Aceh Tengah		1107	Aceh Barat		1201	Nias
	1117	Bener Meriah		1108	Aceh Besar		1202	Mandailing Natal
	1171	Banda Aceh		1109	Pidie		1214	Nias Selatan
	1172	Sabang		1110	Bireun		1216	Pak Pak Bharat
	1173	Langsa		1111	Aceh Utara		1220	Padang Lawas Utara
	1174	Lhokseumawe		1112	Aceh Barat Daya		1224	Nias Utara
	1206	TOBA SAMOSIR		1113	Gayo Lues		1225	Nias Barat
	1211	Karo		1114	Aceh Tamiang		1278	Gunung Sitoli
	1212	Deli Serdang		1116	Aceh Jaya		1301	Kep. Mentawai
	1272	Tanjung Balai		1118	Pidie Jaya		1302	Pesisir Selatan
	1273	Pematang Siantar		1175	Subulussalam		1305	Tanah Datar
	1274	Tebing Tinggi		1203	TAPANULI SELATAN		1306	Padang Pariaman
	1275	Medan		1204	TAPANULI TENGAH		1309	Pasaman
	1276	Binjai		1205	TAPANULI UTARA		1310	Solok Selatan
	1277	Padang Sidempuan		1207	LABUHAN BATU		1311	Dharmasraya
	1374	Padang Panjang		1208	ASAHAN		1377	Pariaman
	1376	Payakumbuh		1209	Simalungun		1402	Indragiri Hulu
	1405	Siak		1210	Dairi		1403	Indragiri Hilir
	1408	Bengkalis		1213	Langkat		1410	Kep. Meranti
	1904	Bangka Tengah		1215	Humbang Hasundutan		1503	Sarolangun
	1906	Belitung Timur		1217	Samosir		1505	Muaro Jambi
	1971	Pangkal Pinang		1218	Serdang Bedagai		1506	Tanjung Jabung Timur
	2101	Karimun		1219	Batu Bara		1507	Tanjung Jabung Barat
	2103	Natuna		1221	Padang Lawas		1508	Tebo
	2171	Batam		1222	Labuan Batu Selatan		1572	Sungai Penuh
	2172	Tanjung Pinang		1223	Labuan Batu Utara		1601	Ogan Komering Ulu
	3171	Jakarta Selatan		1271	Sibolga		1602	Ogan Komering Ilir
	3172	Jakarta Timur		1303	Solok		1603	Muara Enim
	3173	Jakarta Pusat		1304	Sijunjung		1604	Lahat
	3175	Jakarta Utara		1307	Agam		1606	Musi Banyuasin
	3271	Bogor		1308	Limapuluh kota		1607	Banyu Asin
	3272	Sukabumi		1312	Pasaman Barat		1608	OKU Selatan
	3274	Cirebon		1371	Padang		1609	OKU Timur
	3275	Bekasi		1372	Solok		1610	Ogan Ilir
	3276	Depok		1373	Sawah Lunto		1611	Empat Lawang
	3277	Cimahi		1375	Bukittinggi		1703	Bengkulu Utara
	3371	Magelang						

3372	Surakarta	1401	Kuantan Singingi	1704	Kaur
3373	Salatiga	1404	Pelalawan	1705	Seluma
3374	Semarang	1406	Kampar	1709	Bengkulu Tengah
3471	Yogyakarta	1407	Rokan Hulu	1801	Lampung Barat
3506	Kediri	1409	Rokan Hilir	1802	Tanggamus
3510	Banyuwangi	1471	Pekanbaru	1803	Lampung Selatan
3515	Sidoarjo	1473	Dumai	1804	Lampung Timur
3525	Gresik	1501	Kerinci	1805	Lampung Tengah
3572	Blitar	1502	Merangin	1806	Lampung Utara
3573	Malang	1504	Batang Hari	1807	Way Kanan
3574	Probolinggo	1509	Bungo	1808	Tulangbawang
3575	Pasuruan	1571	Jambi	1809	Pesawaran
3576	Mojokerto	1605	Musi Rawas	1810	Pringsewu
3577	Madiun	1671	Palembang	1811	Mesuji
3578	Surabaya	1672	Prabumulih	1812	Tulangbawang Barat
3579	Batu	1673	Pagar Alam	1871	Bandar Lampung
3671	Tangerang	1674	Lubuk Linggau	1872	Metro
3672	Cilegon	1701	Bengkulu Selatan	3202	Sukabumi
3674	Tangerang Selatan	1702	Rejang Lebong	3205	Garut
5101	Jembrana	1706	Mukomuko	3209	Cirebon
5102	Tabanan	1707	Lebong	3303	Purbalingga
5103	Badung	1708	Kepahiang	3305	Kebumen
5171	Denpasar	1771	Bengkulu	3306	Purworejo
6402	Kutai Barat	1901	Bangka	3311	Sukoharjo
6403	Kutai Kartanegara	1902	Belitung	3312	Wonogiri
6405	Berau	1903	Bangka Barat	3313	Karanganyar
6406	Malinau	1905	Bangka Selatan	3314	Sragen
6471	Balikpapan	2102	Bintan	3317	Rembang
6472	Samarinda	2104	Lingga	3318	Pati
6473	Tarakan	2105	Kep. Anambas	3319	Kudus
6474	Bontang	3101	Kep. Seribu	3323	Temanggung
7102	Minahasa	3174	Jakarta Barat	3324	Kendal
7103	Kep. Sangihe	3201	Bogor	3325	Batang
7104	Kep. Talaud	3203	Cianjur	3329	Brebes
7106	Minahasa Utara	3204	Bandung	3403	Gunung Kidul
7108	Siau Tagu Landang Biaro	3206	Tasikmalaya	3501	Pacitan
7109	Minahasa Tenggara	3207	Ciamis	3503	Trenggalek
7111	Bolaang Mongondow Timur	3208	Kuningan	3511	Bondowoso
7171	Manado	3210	Majalengka	3522	Bojonegoro
7172	Bitung	3211	Sumedang	3523	Tuban
7173	Tomohon	3212	Indramayu	3529	Sumenep
7174	Kotamobagu	3213	Subang	3601	Pandeglang
7201	Banggai Kepulauan	3214	Purwakarta	3602	Lebak

7202	Banggai
7502	Gorontalo
7503	Pohuwato
7504	Bone Bolango
7571	Gorontalo
8205	Halmahera Utara
8271	Ternate
8272	Tidore Kepulauan
9101	Fakfak
9102	Kaimana
9171	Sorong
9401	Merauke
9403	Jayapura
9404	Nabire
9411	Puncak Jaya
9412	Mimika
9418	Tolikara
9419	Sarmi
9429	Nduga
9435	Intan Jaya
9471	Jayapura

3215	Karawang
3216	Bekasi
3217	Bandung Barat
3273	Bandung
3278	Tasikmalaya
3279	Banjar
3301	Cilacap
3302	Banyumas
3304	Banjarnegara
3307	Wonosobo
3308	Magelang
3309	Boyolali
3310	Klaten
3315	Grobogan
3316	Blora
3320	Jepara
3321	Demak
3322	Semarang
3326	Pekalongan
3327	Pemalang
3328	Tegal
3375	Pekalongan
3376	Tegal
3401	Kulon Progo
3402	Bantul
3404	Sleman
3502	Ponorogo
3504	Tulungagung
3505	Blitar
3507	Malang
3508	Lumajang
3509	Jember
3512	Situbondo
3513	Probolinggo
3514	Pasuruan
3516	Mojokerto
3517	Jombang
3518	Nganjuk
3519	Madiun
3520	Magetan
3521	Ngawi
3524	Lamongan
3526	Bangkalan

3603	Tangerang
3604	Serang
5107	Karang asem
5201	Lombok Barat
5202	Lombok Tengah
5204	Sumbawa
5205	Dompu
5206	Bima
5207	Sumbawa Barat
5208	Lombok Utara
5272	Bima
5301	Sumba Barat
5302	Sumba Timur
5303	Kupang
5304	TTS
5305	TTU
5306	Belu
5307	Alor
5308	Lembata
5309	Flores Timur
5310	Sikka
5311	Ende
5312	Ngada
5313	Manggarai
5314	Rote Ndao
5315	Manggarai Barat
5316	Sumba Tengah
5317	Sumba Barat Daya
5318	Nagekeo
5319	Manggarai Timur
5320	Sabu Raijua
6101	Sambas
6102	Bengkayang
6103	Landak
6106	Ketapang
6107	Sintang
6108	Kapuas Hulu
6109	Sekadau
6110	Melawi
6111	Kayong Utara
6112	Kubu Raya
6171	Pontianak
6203	Kapuas

		3527	Sampang	6204	Barito Selatan
		3528	Pamekasan	6207	Lamandau
		3571	Kediri	6208	Seruyan
		3673	Serang	6211	Gunung Mas
		5104	Gianyar	6213	Murung Raya
		5105	Klungkung	6304	Barito Kuala
		5106	Bangli	6306	Hulu Sungai Selatan
		5108	Buleleng	6307	Hulu Sungai Tengah
		5203	Lombok Timur	6308	Hulu Sungai Utara
		5271	Mataram	6311	Balangan
		5371	Kupang	7205	Donggala
		6104	Mempawah	7207	Buol
		6105	Sanggau	7303	Bantaeng
		6172	Singkawang	7307	Sinjai
		6201	Kotawaringin Barat	7311	Bone
		6202	Kotawaringin Timur	7313	Wajo
		6205	Barito Utara	7316	Enrekang
		6206	Sukamara	7317	Luwu
		6209	Katingan	7318	Tana Toraja
		6210	Pulang Pisau	7326	Toraja Utara
		6212	Barito Timur	7401	Buton
		6271	Palangkaraya	7402	Muna
		6301	Tanah Laut	7405	Konawe Selatan
		6302	Kota Baru	7409	Buton Utara
		6303	Banjar	7410	Konawe Utara
		6305	Tapin	7603	Mamasa
		6309	Tabalong	7604	Mamuju
		6310	Tanah Bumbu	7605	Mamuju Utara
		6371	Banjarmasin	8105	Kep. Aru
		6372	Banjar Baru	8106	Seram Bagian Barat
		6401	Paser	8107	Seram Bagian Timur
		6404	Kutai Timur	8108	Maluku Barat daya
		6407	Bulangan	8109	Buru Selatan
		6408	Tana Tidung	9103	Teluk Wondama
		6409	Penajem Paser Utara	9104	Teluk Bintuni
		6410	Nunukan	9106	Sorong Selatan
		7101	Bolaang Mongondow	9402	Jayawijaya
		7105	Minahasa Selatan	9415	Asmat
		7107	Bolaang Mongondow Utara	9416	Yahukimo
		7110	Bolaang Mongondow Selatan	9420	Keerom
		7203	Morowali	9426	Waropen
		7204	Poso	9427	Supiori
		7206	Toli-toli	9428	Mamberamo Raya

		7208	Parigi Moutong	9432	Yalimo
		7209	Tojo Una-una	9433	Puncak
		7210	Sigi	9434	Dogiyai
		7271	Palu	9436	Deiyai
		7301	Kep. Selayar		
		7302	Bulukumba		
		7304	Jeneponto		
		7305	Takalar		
		7306	Gowa		
		7308	Maros		
		7309	Pangkajene dan Kep.		
		7310	Barru		
		7312	Soppeng		
		7314	Sidenreng Rappang		
		7315	Pinrang		
		7322	Luwu Utara		
		7325	Luwu Timur		
		7371	Makassar		
		7372	Parepare		
		7373	Palopo		
		7403	Konawe		
		7404	Kolaka		
		7406	Bombana		
		7407	Wakatobi		
		7408	Kolaka Utara		
		7471	Kendari		
		7472	BauBau		
		7501	Boalemo		
		7505	Gorontalo Utara		
		7601	Majene		
		7602	Polewali Mandar		
		8101	Maluku Tenggara Barat		
		8102	Maluku Tenggara		
		8103	Maluku Tengah		
		8104	Buru		
		8171	Ambon		
		8172	Tual		
		8201	Halmahera Barat		
		8202	Halmahera Tengah		
		8203	Kep. Sula		
		8204	Halmahera Selatan		
		8206	Halmahera Timur		
		8207	Pulau Morotai		

		9105	Manokwari		
		9107	Sorong		
		9108	Raja Ampat		
		9109	Tambrauw		
		9110	Maybrat		
		9408	Kep. Yapen		
		9409	Biak Numfor		
		9410	Panjai		
		9413	Boven Digoel		
		9414	Mappi		
		9417	Peg. Bintang		
		9430	Lanny Jaya		
		9431	Mamberamo Tengah		

Lampiran 3

M-files fungsi kernel

```
function K = kernel_func(X1,X2,kernel_,T,p)

[N1 d]      = size(X1);
[N2 d]      = size(X2);

switch kernel_

case 'gauss',
    K = exp(-distSqrd(X1,X2,T));

case 'poly',
    K = (1+X1*T*X2').^p;
end

function D2=distSqrd(X,Y,T)
nx  = size(X,1);
ny  = size(Y,1);

D2  = sum((X.^2)*T,2)*ones(1,ny) + ones(nx,1)*sum((Y.^2)*T,2)' -
2*(X*T*Y')
```

Lampiran 4. Syntax K-Means Clustering

a. M-Files Algoritma K-Means Clustering

```
function [M,j,e] = kmeans(X,K,Max_Its)

%This is a simple and naive implementation of the standard K-Means
%clustering algorithm for the Machine Learning II course.

%The data matrix X (N x D) is passed as argument
%The number of cluster K is passed as argument
%The maximum nos of iterations Max_Its is passed as argument

%The function returns the matrix M (K x D) - corresponding to the
K mean
%vector values

%The functin returns the clusters which each point has been
allocated to
%1.. K in the vector j.

[N,D]=size(X); %N - nos of data points, D dimension of data
I=randperm(N); %a random permutation of integers 1:N - required
               %to set initial mean values

M=X(I(1:K),:); %M is the initial K x D matrix of mean values -
               %simply setting to values of K randomly selected
data
               %points
Mo = M;

for n=1:Max_Its
    %Create distance matrix which is N x K indicating distance
    that each data
    %point is from the current mean values (of which there are K)
    for k=1:K
        Dist(:,k) = sum((X - repmat(M(k,:),N,1)).^2,2);
    end

    %No we simply find which of the K-mean each data point is
    nearest to -
    %so we find the minimum distance of K for each data point.
    This
    %operation can be easily achieved in one line of Matlab using
    the min function.
    [i,j]=min(Dist,[],2);

    %Now that we have the new allocations of points to clusters
    based on
    %the minimum distances obtained form the previous operation we
    can
    %revise our estimates of the position of each mean vector by
    simply
    %taking the mean vlue of all points which have been allocated
    to each
    %cluster using the folowing simple routine.
```

```

for k=1:K
    if size(find(j==k))>0
        M(k,:) = mean(X(find(j==k),:));
    end
end

%we create an N x K dimensional indicator matrix - each row
will have a
%1 in the column corresponding to the cluster that the data
point (row)
%has been allocated to - this is really only required to
compute the
%overall error associated with the current partitioning.

Z = zeros(N,K);
for m=1:N
    Z(m,j(m)) = 1;
end

%This simply prints the current value of the error criterion
which
%K-means is trying to minimise.
e = sum(sum(Z.*Dist)./N);
fprintf('%d Error = %f\n', n, e);
Mo = M;
end

```

b. *Syntax* dengan program R

```

> #Memanggil library yang digunakan
> library(stats)
> kc<-kmeans(as.matrix(data),center=3)
> kc
> kc$cluster
> kc$centers
> kc$totss
> kc$withinss
> kc$tot.withinss
> kc$betweenss
> kc$size

```

Lampiran 5 Syntax Kernel K-Means Clustering

a. M-Files Algoritma Kernel K-Means Clustering

```
function [m, C,z,e] = kernel_kmeans(X,K,Max_Its,kwidth)
%This is a simple implementation of Kernel K-means clustering - an
%interesting paper which proposed kernel based Kmeans clustering
is [1]
%Girolami, M, Mercer Kernel-Based Clustering in Feature Space,
%IEEE Trans Neural Networks, 13(3),780 - 784, 2002.

%Create the kernel matrix.
[N,D]=size(X);
C = kernel_func(X,X,'gauss',kwidth,1);

%initialise the indicator matrix to a random segmentation of the
data
Z = zeros(N,K);
for n = 1:N
    Z(n,rand_int(K)) = 1;
end

%main loop
for its = 1:Max_Its
    %compute the similarity of each data point to each cluster
    mean in
    %feature space - note we do not need to compute store or
    update a mean
    %vector s we are using the kernel-trick - cool eh?
    m=zeros(K,1);
    for k=1:K
        Nk = sum(Z(:,k));
        m(k)=(Z(:,k)'*C*Z(:,k)./(Nk^2)).^0.5;
        Y(:,k) = diag(C) - 2*C*Z(:,k)./Nk +
        Z(:,k)'*C*Z(:,k)./(Nk^2);
    end

    %Now we find the cluster assignment for each point based on
    the minimum
    %distance of the point from the mean centres in feature space
    using the
    %Y matrix of dissimilarities
    [i,j]=min(Y,[],2);

    %this simply updates the indicator matrix Z reflecting the new
    %allocations of data points to clusters
    Z = zeros(N,K);
    for n=1:N
        Z(n,j(n)) = 1;
    end

    %compute the overall error
    e = sum(sum(Z.*Y))./N;
    fprintf('%d Error = %f\n', its, e);
end
```



```

%return the clutsters that each data point has been allocated to
for n=1:N
    z(n) = find(Z(n,:));
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%this is a little utility functino which returns a random integer
between 1
%& Max_Int.
function u = rand_int(Max_Int)
u=ceil(Max_Int*rand);

```

b. *Syntax* dengan program R

```

> #Memanggil library yang digunakan
> library(kernlab)
> data <- read.csv("D:/data_tesis.csv", sep=";", dec=",")
> View(data)
> kernelk=c("rbfdot","stringdot","polydot","tanhdot")
> sc<-kkmeans(as.matrix(data),center=3,kernel = kernelk[1])
> sc
> centers(sc)
> size(sc)
> withinss(sc)
> View(sc)

```

Lampiran 6. Syntax Validitas Kelompok untuk Data Simulasi

```
vals <- vector()
for (k in 2:10) {
  # Perform the kmeans algorithm
  sc <- kkmeans(as.matrix(data), k)

  # Compute the Calinski_Harabasz index
  vals <- c(vals,as.numeric(intCriteria(as.matrix(data),sc," Calinski_Harabasz ")))
}
idx <- bestCriterion(vals," Calinski_Harabasz ")
cat("Best index value is",vals[idx],"\n")

##Kmean
vals <- vector()
for (k in 2:10) {
  # Perform the kmeans algorithm
  kc <- kmeans(as.matrix(data), k)

  # Compute the Calinski_Harabasz index
  vals <-
c(vals,as.numeric(intCriteria(as.matrix(data),kc$cluster,"Calinski_Harabasz")))
}
idx <- bestCriterion(vals," Calinski_Harabasz ")
cat("Best index value is",vals[idx],"\n")
```

Lampiran 7. Syntax Validitas Kelompok untuk Data Penelitian

```
vals <- vector()
for (k in 10) {
  # Perform the kmeans algorithm
  sc <- kkmeans(as.matrix(data), 10)

  # Compute the Calinski_Harabasz index
  vals <- c(vals,as.numeric(intCriteria(as.matrix(data),sc,"Gamma")))
}
idx <- bestCriterion(vals,"Gamma")
cat("Best index value is",vals[idx],"\n")
##Kmean
vals <- vector()
for (k in 10) {
  # Perform the kmeans algorithm
  kc <- kmeans(as.matrix(data), 10)

  # Compute the Calinski_Harabasz index
  vals <- c(vals,as.numeric(intCriteria(as.matrix(data),kc$cluster,"Gamma")))
}
idx <- bestCriterion(vals,"Gamma")
cat("Best index value is",vals[idx],"\n")
```

Lampiran 8. Syntax Untuk Membangkitkan Data Simulasi

a. Data Linier

```
linier=function(n)
{
  rasio=-1
  while (rasio < -0.1| rasio >0.1)
  {
    x1=rnorm(n,0,1); x2=rnorm(n,0,1)
    y=matrix(ncol=1,nrow=n)
    for (i in 1:n)
    {
      r=x1[i]+x2[i]
      if (x1[i]>x2[i])
        y[i]=1 else y[i]=-1 }
    rasio=sum(y)/n
  }
  linier=cbind(x1,x2,y)
  write.csv(linier,'d:/linear.csv')
}
```

b. Data Non Linier (Lingkaran)

```
lingkaran=function(n)
{
  rasio=-1
  while (rasio < -0.1| rasio >0.1)
  {
    x1=rnorm(n,0,1) ; x2=rnorm(n,0,1)
    y=matrix(nrow=n,ncol=1)
    r=mat.or.vec(4,1)
    for (i in 1:n)
    {
      r1=sqrt(((x1[i])^2+(x2[i])^2))
      if (r1< 1.2 ) y[i]=1 else y[i]=-1 }
    rasio=sum(y)/n
    print(rasio)
  }
  lingkaran=cbind(x1,x2,y)
  write.csv(lingkaran,'d:/lingkaran.csv')
}
```

BIOGRAFI PENULIS



Penulis dilahirkan di Bantul, Yogyakarta pada tanggal 11 Maret 1983, merupakan putri keempat dari tujuh bersaudara, dari pasangan Bapak Ngadiman dan Ibu Mujiyem. Penulis telah menempuh pendidikan formal yaitu di TK PKK Pamardisiwi (1988-1989), SDN Jejeran II (1989-1995), SLTP Negeri 10 Yogyakarta (1995-1998), SMU Negeri 8 Yogyakarta (1998-2001). Kemudian penulis melanjutkan pendidikan ke jenjang sarjana di Sekolah Tinggi Ilmu Statistik (STIS) Jakarta (2001-2006) jurusan Statistik peminatan Ekonomi. Setelah menyelesaikan pendidikan DIV di STIS, penulis ditugaskan bekerja di BPS Kabupaten Sumba Barat Provinsi Nusa Tenggara Timur sebagai staf seksi Statistik Neraca Wilayah dan Analisis Statistik. Pada tahun 2011 penulis pindah tugas ke BPS Kabupaten Kulon Progo Provinsi D.I. Yogyakarta. Pada tahun 2013 penulis memperoleh kesempatan untuk mendapatkan beasiswa dari BPS untuk melanjutkan jenjang pendidikan S2 di Jurusan Statistika Fakultas MIPA Institut Teknologi Sepuluh Nopember Surabaya. Pembaca yang ingin memberikan kritik, saran dan pertanyaan mengenai penelitian ini, dapat menghubungi penulis melalui email maysaroh.bps@gmail.com.